



Mathematics/ Science/ Technology
An Inter-faculty Second Level Course
Mathematical Models and Methods

mathematical models and methods

Unit M

Mathematical Modelling

Mathematics/Science/Technology
An Inter-faculty Second Level Course
MST204 Mathematical Models and Methods

Unit M

Mathematical modelling

Prepared for the Course Team
by Mike Crampin

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 1993. Reprinted 1993, 1995.

Copyright © 1993 The Open University.

All rights reserved. No part of this work may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, without written permission from the publisher or a licence from the Copyright Licensing Agency Limited. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd of 90 Tottenham Court Road, London, W1P 9HE.

Edited, designed and typeset by the Open University using the Open University T_EX System.

Printed and bound in the United Kingdom by Staples Printers Rochester Limited, Neptune Close, Medway City Estate, Frindsbury, Rochester, Kent ME2 4LT.

ISBN 0 7492 2191 7

This text forms part of an Open University Second Level Course. If you would like a copy of *Studying with The Open University*, please write to the Central Enquiry Service, PO Box 200, The Open University, Walton Hall, Milton Keynes, MK7 6YZ. If you have not already enrolled on the Course and would like to buy this or other Open University material, please write to Open University Educational Enterprises Ltd, 12 Cofferidge Close, Stony Stratford, Milton Keynes, MK11 1BY, United Kingdom.

Contents

Introduction	5
1 First thoughts about the modelling process	7
Summary of Section 1	13
2 Two models for you to analyse	14
2.1 Modelling drug therapy	14
2.2 Skid marks	19
Summary of Section 2	21
3 The skills of modelling	22
3.1 Interpreting the question	22
3.2 Choosing appropriate variables	23
3.3 Simplifying the problem and identifying its most important features	24
3.4 Specifying relationships	24
3.5 Interpreting and testing the results	26
3.6 Improving the model	27
Summary of Section 3	27
4 Putting it all together	28
Summary of Section 4	37
5 A modelling exercise for you to tackle	38
Appendix: Solutions to the exercises	39

Introduction

As you know, one of the main aims of this course is to help you to develop your appreciation of the role of mathematics in understanding and predicting the behaviour of the real world (as opposed to the world of mathematical theories), and to begin to acquire the skill of using mathematics to help to solve real-world problems, which we call mathematical modelling. This unit is devoted to that task.

There was a time when the application of mathematics was restricted more or less to physics and engineering; but those days are long gone, and now mathematics turns up all over the place. In case you need to be convinced that this is the case, you may like to consider the following list of titles of conferences organized in the course of a year by the Institute of Mathematics and its Applications.

Mathematics in Industrial Maintenance

(topics included component replacement problems, availability and downtime modelling, and inventory and manpower problems)

Business Modelling

(themes: the mathematical modelling of production and manufacturing, finance, human resource management, and marketing)

The Mathematical Theory of the Dynamics of Biological Systems

(which covered the epidemiology of infectious diseases, including AIDS, as well as the harvesting, control and regulation of biological populations)

Control Theory

(economic systems and neural networks were among the topics here)

The Mathematics of Surfaces

(which was concerned with the representation of complicated surfaces, and dealt with computer-aided design and machining, robot vision, cartography, and even applications in television and the film industry)

Stably Stratified Flows

(it may not be apparent from the title, but this conference was actually about the effects of topography on the weather)

Mathematics in Food Production, Processing and Preservation

(covering such topics as freezing and chilling, cooking, and extrusion and injection)

Aerospace Vehicle Dynamics and Control

(everything from airships to spacecraft)

Mathematics in Signal Processing

(with applications to astronomy, communications, medicine, meteorology, radar, etc.)

Puzzles in Quantum Theory

(physics has not been *completely* overwhelmed by newer areas of application)

All human life is here! Though these diverse uses of mathematics call on different aspects of the subject, there is a range of skills which is common to all of them, and these are the skills that together constitute the craft of mathematical modelling.

This unit takes an incremental approach to the teaching of mathematical modelling. In the first section you will be introduced to the art of interpreting and analysing a model produced by someone else, and in the second you will be asked to carry out this process for yourself. In the third section you will practise some of the individual skills of modelling. The fourth and fifth sections are devoted to integrating those skills.

Mathematical modelling can be learnt only by doing it, so there are many exercises in this unit: in particular, in Section 5 you are asked to develop a simple mathematical model yourself. Sections 1 and 4, however, consist mostly of reading; but I do hope that you will read them actively, rather than passively – by which I mean that I hope you will try to answer for yourself the many questions that are implicit in the text, and try to anticipate the direction that the argument will take.

THE SUNDAY TIMES magazine

The Transport and Road Research Laboratory has devised
a formula to enable you to calculate your
mathematical chance of having a traffic accident next year.
The formula you have to apply is this:

$$A_c = 0.00633 \exp \{s + g\} \\ (1 + 1.6p_d) \\ (p_b + 0.65p_r + 0.88p_m) \\ M^{0.279} \\ \exp \{b_1/Ag + b_2/(X + 2.6)\}$$

a year of glorious follies

Two television programmes on modelling will be broadcast during the two weeks that are set aside in the course calendar for your work on this unit and the associated assignment. Each of these programmes will further illuminate the ideas discussed in the unit. Neither requires you to do any specific work either before or after the broadcast, and unlike the other programmes for the course, neither introduces mathematical results which you will need to refer to later. I have therefore not thought it necessary to include programme summaries in the unit.

1 First thoughts about the modelling process

Some years ago, *The Sunday Times* magazine for the last Sunday of publication of the year appeared with the cover shown here on page 6. Among the other 'glorious follies' to be found inside were photographs of the Prime Minister with his trousers held up by a paper clip, Prince William falling off a horse, Benetton adverts, and sumo wrestlers. On page 3 of the magazine the reader was told that 'the probability of having a road accident (see cover) is only slightly more difficult to calculate than the running total of Liz Taylor's husbands'.

Reading the magazine sections of Sunday newspapers might very well be regarded as folly, and not very glorious at that, so perhaps one shouldn't be too surprised when mathematical modelling is made into a joke so prominently in one of them. I don't know how you react, but it took me several minutes to recover from my annoyance that a so-called quality newspaper should take such a dismissive attitude to what is clearly a serious investigation of a serious subject. I suppose that I should have learnt over the years that mathematics is rarely taken seriously by the media, even though it plays such an important role in modern life.

Once I had got over my irritation, however, I began to think about the formula itself, to ponder its significance, and to try to understand what it means. Many of the points that I want to raise in this unit came up during my consideration of the formula, so I'd like to recount some of my thoughts by way of introduction to the unit. By describing my reflections about the formula, I can bring to your attention again some of the points about modelling that were first raised (in this course) in *Unit 3*, and in particular to say once more what mathematical modelling involves and why it is important. I should make it clear from the outset that the model we are about to consider has been produced by professionals at the Transport and Road Research Laboratory, and is more complicated and sophisticated than any model you will be asked to develop yourself in this course; but I don't think you will find it difficult to appreciate the main issues it raises.

This section could be described as an exercise in interpreting a mathematical model: not an easy task in this case, since there is so little information to go on. There is not a great deal for you to do except read; but you will probably find it valuable to think about how the points that arise in the discussion apply to other models that you have met, earlier in this course or elsewhere. In the next section you will find some extensive exercises in which you will be asked to analyse and interpret some mathematical models yourself.

The formula we have to deal with is this:

$$A_c = 0.00633 \exp\{s + g\}(1 + 1.6p_d)(p_b + 0.65p_r + 0.88p_m)M^{0.279} \exp\{b_1/Ag + b_2/(X + 2.6)\};$$

it is a mathematical model of the frequency of traffic accidents. Here already is my first point about modelling: a mathematical model is a formula. Perhaps that is a little too sweeping, but it does express a basic truth: the aim of mathematical modelling is to capture the essentials of some situation or process in the form of one or more equations or other mathematical relationships, in order to understand it better, to predict how it may develop, or to help one make decisions.

The Sunday Times dismissed the formula as a 'glorious folly', implying perhaps that it has no use. We may reserve judgement on that, but one thing is certain: there is no way of telling how useful it may be without being able to say what it means, and *The Sunday Times* has made this task more or less impossible by not revealing what the symbols stand for. You and I can see the general structure of the formula: the right-hand side consists of a product of several terms, some of which are linear functions of certain variables ($(1 + 1.6p_d)$, for example), some of which involve exponentials, and one of which is a power. The formula 'enables you to calculate your mathematical chance of having a traffic accident next year'. (Why 'mathematical' chance, I wonder – what is an unmathematical chance?) So A_c is presumably the chance of an individual having a traffic accident during one year. I have deliberately changed the wording here: *The Sunday Times* seems to suggest that the formula is personal to you (the reader), and applies only to the year following publication, which is hardly likely to be the case. In order for *you* to calculate *your* chance of having an accident next year, you must substitute those values of the variables s , g , p_d and so on which are appropriate to you in the year in question. We aren't told what kind of accident, incidentally; my guess is that the formula is concerned with motor accidents, if only because accidents to pedestrians or cyclists generally seem to be of much less interest to journalists. It also seems likely that the formula deals with drivers rather than passengers, and with cars rather than with other vehicles.

So the formula expresses how an individual's chance of having a car accident depends on certain other variables; but what those variables are one cannot tell just by gazing at the front cover of *The Sunday Times* magazine. I became sufficiently interested in the question to write to the Transport and Road Research Laboratory for more information. But inquisitiveness won't wait for the post (and in any case Christmas and the New Year had to pass before I could expect a reply); so I found myself, in the meantime, thinking about what factors do affect one's chances of having a road accident. I found it a quite instructive exercise to try to list the things about an individual's motoring which would be likely to influence the probability of having an accident, and to ask how, qualitatively, they might affect the answer. For example, how much one uses a car is obviously important: the more one drives, the more likely one is to be involved in an accident. That is worded rather loosely, but the idea could easily be made sufficiently precise to make it worthy of representation by a symbol in an equation: annual mileage (or, if you prefer, annual kilometrage). You may wish to stop reading at the end of this paragraph and spend a few minutes trying to extend the list for yourself. It is probably better not to try to guess what the symbols in the formula mean, but to think about the problem from scratch.

I listed the following features, in addition to annual mileage.

- The driver's age and state of health.
- The driver's skill, tendency to drive fast, etc.
- The type of road encountered (it is often claimed that motorways are safer than other roads, for example).
- The age and condition of the car.
- The driver's gender.

I don't claim that this is an exhaustive list, but I do think that all these are relevant factors. Whether they could all be defined in such a way that they could appear in a formula like the one under consideration is a matter that requires further thought. One way of approaching this question is to ask whether and how each factor can be quantified. Age clearly can; state of health, and skill, are not so easy to measure directly. On average, one's state of health tends to decline with age (alas), so in the interest of simplification it might be as well to concentrate on the single variable age. Likewise, to include skill I would have to think of some other, more easily measurable, factor to which it is closely related. Gender may appear to cause difficulties: it does strain the English language a bit to describe someone's gender as a variable. Remember, however, that the purpose of the exercise is to obtain a general formula which can be made to apply to any individual by giving the symbols (or variables) the appropriate values. It isn't too difficult to see how this might work for the person's

gender: one merely has to allow the corresponding symbol to take only two values, say 0 for a man and 1 for a woman.

My second point about modelling is this. If the primary aim, in modelling, is to describe a situation in terms of an equation or similar mathematical relationship, then it is clearly of the greatest importance to decide what variables are going to enter the equation. Very often (as here) one will want to investigate one particular variable – in this case, the chance of having an accident. It is then sensible to ask oneself what factors it might depend on. The answers must, of course, be capable of clear definition and measurement, since they are to appear in the equation as variables. One should keep at the back of one's mind the possibility that a variable may have only a limited range of allowable values: perhaps as few as two. Incidentally, the variable A_c is called the *dependent* variable, because it depends on the others; the other variables are called *independent* variables.

When I'm thinking about the variables that enter a mathematical model, I make a practice of trying to see how, qualitatively, each will affect the outcome. For example, I think it is pretty well bound to be the case that older and more skilful drivers are likely to be safer. I could express this mathematically by saying that the chance of having an accident decreases with increasing age, all other factors being equal; or that the chance of having an accident is a decreasing function of age. Another way of representing the dependence of the chance of having an accident on age would be to use a graph: I hope you can see (in your mind's eye) what such a graph must look like, in general terms.

Here is my third point – a hint rather than an instruction: you can gain some useful initial insight into a model by thinking about the qualitative nature of the relationships between the variables.

So far, I have been describing my speculations while I was waiting to hear from the Transport and Road Research Laboratory (TRRL). I did eventually receive a reply to my enquiry, and found that I was correct in assuming that the report is concerned with accidents to drivers of cars. I discovered that the symbols in the formula actually have the following meanings.

G. Maycock, C. R. Lockwood and Julia F. Lester, 'The accident liability of car drivers', TRRL Research Report 315 (HMSO, 1991).

- A_c is the annual accident frequency.
- s and b_2 are variables whose values depend on the driver's sex:
 - $s = 0, b_2 = 3.5$ for males;
 - $s = -0.02, b_2 = 2.3$ for females.
- g and b_1 are variables whose values depend on the driver's socio-economic group:
 - $g = 0, b_1 = 13$ for drivers in socio-economic groups A, B and C1;
 - $g = -0.72, b_1 = 23$ for drivers in socio-economic groups C2, D and E.
- p_d is the proportion of driving undertaken in the dark.
- p_b, p_r and p_m are the proportions of driving undertaken in built-up areas, rural areas and on motorways, respectively.
- M is the distance driven annually, in miles.
- Ag is the age of the driver, in years, at the mid-point of the accident period.
- X is the number of years since passing the driving test, also determined from the mid-point of the accident period.

I found much of this reassuring, and the rest interesting. The authors of the TRRL report don't talk of chance, but of frequency, of accidents. The idea is that for every hundred people (say) of the same age and gender, driving the same number of miles, and so on, there will be $100A_c$ accidents in a year. Gender comes into the formula in quite a complicated way: it both modifies the overall constant 0.00633, via the factor e^s , and has an effect on the way that the accident frequency depends on the number of years since passing the driving test. It appears that class (more technically, socio-economic group) is a factor, which works in a way similar to gender; I hadn't anticipated this. (Incidentally, the socio-economic groups are defined by occupation, as follows: Group A consists of the professional occupations (such as mechanical engineer); Group B of intermediate occupations (garage proprietor); Group C1 of

non-manual skilled occupations (driving instructor); Group C2 of manual skilled occupations (motor mechanic); Group D of partly skilled occupations (traffic warden); and Group E of unskilled occupations (car-park attendant).)

Another factor I missed is the proportion of driving done in darkness, which on second thoughts is clearly important. I find the classification of road types a bit strange, though the principle seems clear enough: I think I would have chosen two separate classifications, into grade of road (motorway, A class, B class) and situation (urban or rural). Note that the variables p_b , p_r and p_m are not independent: since they are proportions, and (by implication) one's driving must always fall into one or other of the corresponding categories, they must satisfy $p_b + p_r + p_m = 1$.

The TRRL report is rather careful to state the units in which variables are measured, where relevant: distances, you will note, are measured in miles. This is an important thing to remember when defining variables.

The time variables Ag and X are also specified very carefully: for example, Ag is not just described as the driver's age, but the driver's age 'in years, at the mid-point of the accident period'. It is worth noting in passing that there are limits on the values these variables can take: for example, Ag must be a whole number and cannot be less than 17; there is also an upper limit to its possible values, though this is not so sharply determined – perhaps 100 would be a reasonably safe choice. It is good practice to look out for limitations on values like this. I suppose that X is to be regarded as a measure of experience, and to some extent may be considered to stand for skill, which (as I remarked earlier) is difficult to measure directly.

One point about the symbols used: Ag is a single variable, it does not mean $A \times g$. I would never use such a symbol myself, because of the risk of confusion. Choosing symbols well is a minor but useful art. It helps if each symbol carries some reminder of the thing it symbolizes, and it is therefore natural to use the first letter of an appropriate word for the symbol. In this case this approach leads to a difficulty because two of the main variables, accident frequency and age of driver, begin with the letter 'a'. But I think that the use of a or α would have been a better solution to this difficulty than the one actually adopted.

One factor that I raised is omitted from the formula, namely the age and condition of the car. I remain puzzled by the absence of any variable relating to the car. This is not to say that the TRRL model is wrong: it is as well to be aware from the start that mathematical modelling (unlike solving quadratic equations, for example) is not the sort of activity in which there is a single correct answer to every problem. In many cases there is room for different approaches; this adds to both the interest and the difficulty of modelling (in my view, anyway).

Now that we know what the symbols in the formula mean, we can think about some of its implications. Let us look first at the effect of mileage. If all the other variables are held fixed, so that we consider the effect of mileage alone, we see that A_c is proportional to $M^{0.279}$, in other words $A_c = kM^{0.279}$ for some k independent of M . Now the graph of $y = x^{0.279}$ is similar to that of $y = x^{0.25}$, or $y^4 = x$, which looks like Figure 1.

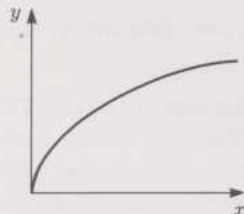


Figure 1

So, as one would expect, the accident frequency increases with increasing mileage – and is 0 if $M = 0$, which is as it should be! But it is interesting to see just how the accident frequency increases. One might have expected that doubling the mileage would double the frequency; but the formula predicts that the frequency would be multiplied by $2^{0.279} \simeq 1.2$, that is, that it would be increased by only about 20%.

The formula $A_c = kM^{0.279}$ can itself be regarded as a simple model of the frequency of road accidents; indeed, it is a perfectly satisfactory model, so long as one is interested only in how the frequency depends on mileage. The factor k contains information about driving in the dark, the driver's age, and so on; we can tease this information out, at the price of making the model progressively more complicated. It is a well-established principle of mathematical modelling that it is best to start with as simple a model as possible and to make it progressively more complicated by incorporating additional features.

Turning our attention to the effect of driving in the dark, if the proportion of driving done in darkness is increased, then the likelihood of an accident increases, all other things being equal. If all driving is done in darkness ($p_d = 1$), then the accident frequency is 2.6 times greater than if all driving is done in daylight ($p_d = 0$). At a guess, I'd say that in the depths of winter about a third of my driving is done in the dark, whereas at the height of summer practically none of it is: according to the formula I'm about 50% more likely to have an accident during the winter, which is a sobering thought.

As I had suspected, motorway driving is safer than driving in built-up areas: this follows from the fact that the coefficient of p_m is less than the coefficient of p_b ; but driving on rural roads is safer than both of these.

Increases in age and in experience both decrease accident frequency, since both Ag and X appear in reciprocals. However, the effect of either depends on the gender or class of the driver. Why gender is paired with experience, and age with class, is not at all clear to me. However, let us look at gender and experience. When all other variables are fixed, we have

$$A_c = l \exp\{s\} \exp\{b_2/(X + 2.6)\} = l \exp\{s + b_2/(X + 2.6)\},$$

where l stands for the product of all the other factors in the formula. For males $s = 0$ and $b_2 = 3.5$, while for females $s = -0.02$ and $b_2 = 2.3$. It follows that, for a given amount of experience, accident frequency is less for a woman than for a man. How much less depends on the level of experience. For completely inexperienced drivers, for whom $X = 0$, the ratio of the frequencies (woman to man) is

$$\exp\{-0.02 + 2.3/2.6\} : \exp\{3.5/2.6\} \simeq 2.37 : 3.84 \simeq 0.6 : 1,$$

so the accident frequency for inexperienced female drivers is only about 60% of that for inexperienced male drivers of the same age, annual mileage and so on. For drivers with 20 years' experience, however, the difference is much less marked, the ratio being

$$\exp\{-0.02 + 2.3/22.6\} : \exp\{3.5/22.6\} \simeq 1.09 : 1.17,$$

i.e. the accident frequency is almost the same for women and men. In fact, the pure gender factor e^s is relatively insignificant ($e^{-0.02} \simeq 0.98$), the main difference coming from the experience term; as X increases, this difference rapidly becomes less marked, as the graphs in Figure 2 show.

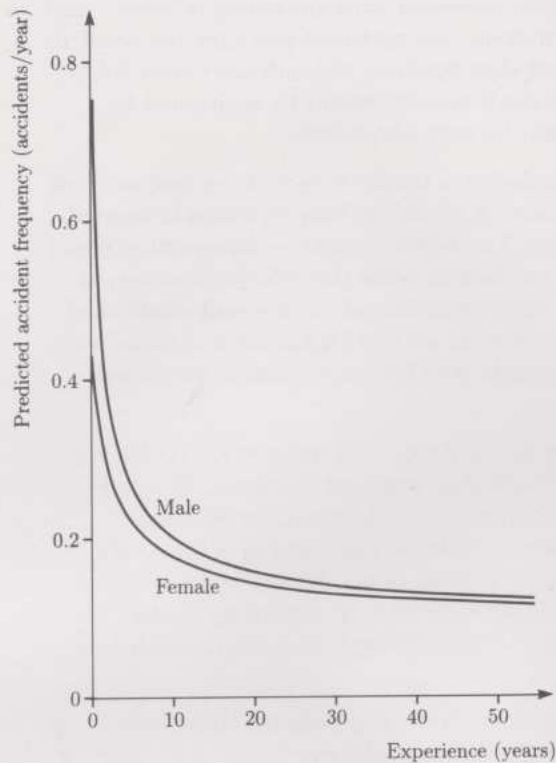


Figure 2

These graphs are drawn for the same fixed value of l .

I have described several conclusions that can be drawn from the model, and there are of course many more. Some of these conclusions merely confirmed expectations, while others were in some way new and surprising. The general point to note here is that mathematical modelling does not end when one has constructed a model: one has to investigate and interpret its implications, both to confirm its validity and to draw new conclusions.

It is time to return to the question of whether all this is a glorious folly. The view expressed in *The Sunday Times*, that it is, presumably rests on two assumptions. One of them is stated quite explicitly on page 3 of the magazine: anything that involves calculations more complicated than counting up to eight is regarded by the newspaper's editors as beyond the capabilities of most of its readers. But there also seems to be an implicit assumption that the exercise of trying to analyse the incidence of road accidents and the factors that affect their frequency is intrinsically pointless. That seems to be the view of the journalists on *The Sunday Times*; the staff of the Transport and Road Research Laboratory presumably thought otherwise.

In examining the implications of the formula, I have shown that it reproduces some well-known or obvious facts (if that is the word) about safety in driving: motorways are safer than urban roads, experienced drivers are involved in fewer accidents than are inexperienced ones, the more you drive the more likely you are to have an accident. It is reassuring that the formula does confirm one's expectations in this way – if it did not, one would have little faith in it (though it is as well to remember that 'well-known facts' do sometimes turn out to be myths). If the formula was nothing more than a complicated way of restating the obvious, then deriving it would indeed have been a somewhat pointless exercise. But surely it does contain new information. I was surprised by a number of the conclusions drawn above: the effect of darkness, the relative safety of rural roads, the way the significance of mileage tails off as mileage increases, the fact that experience irons out the difference between the sexes.

I think that the very personal way in which the formula was presented in *The Sunday Times* ('a formula to enable you to calculate your mathematical chance of having a traffic accident') both misrepresents it and undervalues it. Though there are some pointers to ways in which you and I can reduce our chances of having an accident (try to avoid driving in the dark, use motorways and rural roads in preference to urban

ones), they are few in comparison with things over which one has no control. But it is scarcely likely that the TRRL intended the formula to be used like this. In fact, the formula was obtained as part of a continuing programme of research 'intended to provide a factual basis for more detailed psychological studies' of the 'relationship between the accident liability of an individual driver . . . and factors such as age, sex and driving experience'. The TRRL report on this research describes its aims as follows.

The effect of age and experience on accident liability are factors of particular interest since there are grounds for believing that they influence driving behaviour (and hence accidents) differently. Age effects clearly reflect growing maturity – perhaps changing perceptions of risk, or a growing sense of social responsibility. Experience on the other hand reflects a learning process within the driving task. The distinction between maturation and learning could have important implications for the application of road safety remedial treatments. If age is the dominant factor, then short of devising a method of accelerating the maturation process, raising the driving age might be the only practical countermeasure. If on the other hand, experience plays a crucial part, then improved safety might be achieved by devising better ways of imparting those skills necessary for safe driving to novice drivers – a matter of training.

Other possible uses of the results suggest themselves: for example, the analysis of the effect of driving in the dark on accident frequency should be useful in the context of the debate about summer and winter time, and knowledge of how age affects safety should be of interest to insurance companies in determining how premiums should be fixed for drivers of different ages. While this may not be the most far-reaching piece of research on road safety that has ever been done, to call it a glorious folly does seem to me to be a bit unfair.

It is worth noting that an analysis of this kind itself generates further questions for investigation and ideas for improvement. Here, as the authors of the TRRL report point out, the nature of the dependence of accident frequency on mileage is crying out for explanation. Furthermore, the analysis could be improved by finding a variable to stand for 'experience' better than the number of years since passing the test. It rarely turns out to be the case that there is a final word in mathematical modelling: models are almost always capable of extension and improvement.

I am conscious that nowhere in this section, so far, have I explained how the TRRL formula was actually derived. It was actually done by collecting a lot of data about accidents and fitting curves to them, a process related to that described in Subsection 1.2 of *Unit 3* under the name of empiricism. Here, though, the problem is rather more complicated, since there are several variables involved rather than just one, and the techniques involved are beyond the scope of this course. The formulation of a model is clearly the fundamental step in mathematical modelling, and we must get to grips with it eventually; but it is a good idea if first you become more familiar with the process of modelling as a whole, and that is the aim of both this and the next section.

Summary of Section 1

A mathematical model of a real-world situation or process is an equation, a set of equations, or some other mathematical relationship between variables which specify the possible states of the system being modelled. In order to understand and analyse a model, therefore, it is necessary to know how the variables are defined, and their units. Interpreting the model involves working out what it states about the relationships between the variables; a successful model will both confirm what one knows already about the system under consideration, and predict new conclusions. Models are usually developed with a specific purpose in mind, often to solve some particular problem relating to the real-world situation being modelled; being aware of this purpose is clearly also an important aspect of interpreting the model.

2 Two models for you to analyse

In the previous section I made some general points about the processes involved in mathematical modelling, in the context of a description of a specific model concerned with the frequency of road accidents. I am now going to ask you to confirm your understanding of those points by analysing and interpreting accounts of two mathematical models. There are two long exercises for you to do in this section: each consists of a number of questions about the model described in the preceding report. I have selected my examples from prerequisite courses, so you may find one of them familiar, but not so familiar that you will not have to think carefully in order to answer the questions.

The first model is taken from *M101* and *MS284*. While you are reading it, bear in mind the points made in the previous section, especially the summary. Look out for the purpose of the model, the process which is being modelled, the definitions of the variables, the derivation of the relation between them, the conclusions that are drawn from the model, and any possibilities for improving and developing the model.

2.1 Modelling drug therapy

In this report, drug quantities are measured in milligrams and volumes in litres. We defined these units when introducing the SI system, but you may like to be reminded that

$$1 \text{ milligram (mg)} = 10^{-3} \text{ grams (g)} = 10^{-6} \text{ kilograms (kg)}$$

and that

$$1 \text{ litre (l)} = 10^3 \text{ cubic centimetres (cm}^3\text{)} = 10^{-3} \text{ cubic metres (m}^3\text{)}.$$

Patients with asthma suffer from constriction of the airways in the lungs and, consequently, difficulty in breathing out. This ailment can be alleviated by introducing the drug theophylline into the bloodstream. This is done by injecting another drug, aminophylline, which the body quickly converts to theophylline. Once present in the blood, however, the drug is steadily excreted from the body via the kidneys. In other words, the system 'leaks', and unless there is replenishment, the quantity of drug in the blood will fall.

It is known from experiment that theophylline has hardly any therapeutic effect if its concentration in the bloodstream is below 5 mg/l (milligrams per litre), and that concentrations of above 20 mg/l are likely to have toxic effects. The problem is to administer the drug in such a way that the concentration remains within the therapeutic range between 5 mg/l and 20 mg/l.

Experimental measurements can be obtained by injecting a known dose of aminophylline into a patient, allowing time for it to diffuse throughout the bloodstream and subsequently finding the concentration of theophylline in blood samples taken at regular intervals. One then knows the initial quantity put into the bloodstream and the concentration at various later times, but there is no time for which both the quantity and concentration are known. Figure 1 shows a graph of concentration against time which could be obtained from such an experiment.

Most of this subsection (up to Exercise 1) is taken from *M101 Block V Unit 3*, Section 3.4, or *MS284 Unit 16*, Section 4.

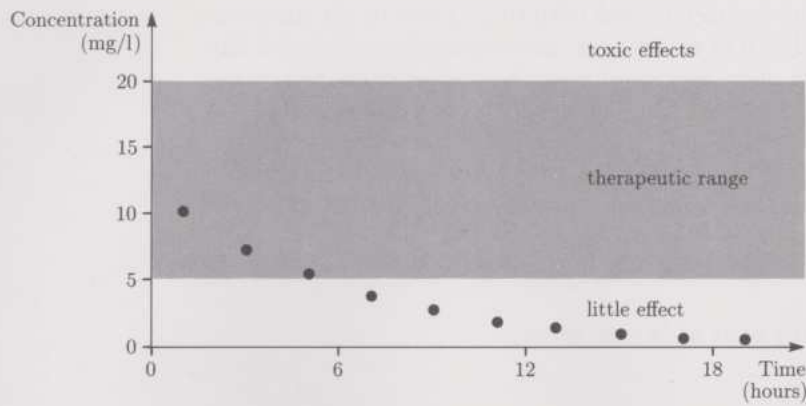


Figure 1

The model is based on the following assumptions.

- The conversion from aminophylline to theophylline takes place at the moment of injection. This is equivalent to assuming that it is theophylline itself which is injected.
- The theophylline is distributed uniformly throughout the bloodstream upon injection and at all subsequent times. Hence the drug's concentration at a given time is the same at all points of the bloodstream.
- The quantity of theophylline in the body is proportional to its concentration in the blood. The constant of proportionality is called the *apparent volume of distribution*. This is not simply the volume of blood in the body, as the drug is also present in other bodily fluids and tissues. It can be thought of as the total volume of blood for which the drug concentration would be equal to that in the actual bloodstream.
- No theophylline enters the body between injections.
- The quantity of drug excreted from the body in a given time interval is proportional both to the total quantity present at the start of the interval and to the duration of the interval.

On introducing the variables t (time in hours), q (quantity of theophylline present in mg) and c (drug concentration in the blood in mg/l), we have $q = q(t)$ and $c = c(t)$.

Assumption (c) can be expressed as

$$c(t) = \frac{q(t)}{V},$$

where V is the apparent volume of distribution in litres.

Assumptions (d) and (e) are used in concert with the input-output principle, which is

$$\boxed{\text{accumulation}} = \boxed{\text{input}} \text{ minus } \boxed{\text{output}}.$$

The principle is applied to the quantity of theophylline in the body for an interval between times t and $t + \delta t$ during which no injection is administered. The accumulation of the drug over this interval is equal to the difference between the quantities present at the end and at the start of the interval, that is,

$$\text{accumulation} = q(t + \delta t) - q(t).$$

By assumption (d) there is no replenishment of the drug during the interval, and so

$$\text{input} = 0.$$

Assumption (e) states that the output is proportional both to $q(t)$ and to the duration δt of the interval. In fact, the second half of this is only approximately correct, but the approximation becomes progressively more accurate for smaller values of δt . Expressing assumption (e) in symbols, we have

$$\text{output} \simeq \lambda q(t) \delta t,$$

where λ is a positive constant. The input-output principle, therefore, leads to the equation

$$q(t + \delta t) - q(t) \simeq -\lambda q(t) \delta t.$$

Dividing through by δt and taking the limit as $\delta t \rightarrow 0$ gives

$$\frac{dq}{dt} = -\lambda q.$$

This differential equation for $q(t)$ can be supplemented by an initial condition. If a dose D mg is injected at time $t = 0$ into a patient who was previously free of the drug then, according to assumption (a), we obtain

$$q(0) = D.$$

Given the differential equation

$$\frac{dq}{dt} = -\lambda q \quad (t > 0)$$

and the initial condition

$$q(0) = D,$$

we must find an expression for $q(t)$, then use the relation

$$c(t) = \frac{q(t)}{V}$$

to deduce an expression for $c(t)$. The general solution is

$$q = Ae^{-\lambda t},$$

where A is an arbitrary constant, but since $q > 0$ in the current model we are concerned here only with positive values of A .

Putting $t = 0$ into the general solution and using the initial condition $q(0) = D$ gives

$$D = A,$$

so that the required particular solution is

$$q = De^{-\lambda t}.$$

Then we have

$$c = \frac{D}{V} e^{-\lambda t},$$

where D/V is the initial concentration.

A qualitative prediction of the exponential decay solution is that between injections the concentration will fall with time but more and more slowly.

The form of the solution is itself a quantitative prediction, but it can be put into a form more amenable to comparison with reality. Taking the logarithms of both sides of the equation for c gives

$$\log_e c = \log_e \left(\frac{D}{V} \right) - \lambda t.$$

This predicts that a graph of $\log_e c$ against t should be a straight line.

The qualitative prediction is satisfied by the experimental results shown in Figure 1 (page 15). The quantitative prediction is also satisfied, for when the experimental data are plotted as suggested the result is a straight line (see Figure 2).

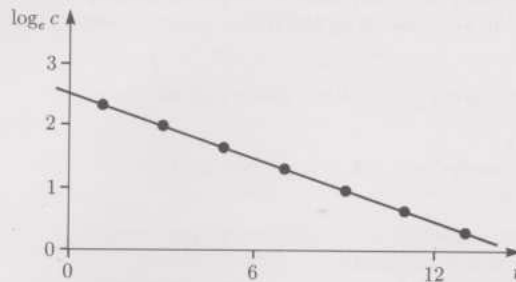


Figure 2

To two significant figures, the slope of this line is -0.17 and the intercept on the vertical axis is 2.5 . Thus we have

$$\lambda = 0.17,$$

and

$$\log_e \left(\frac{D}{V} \right) = 2.5$$

or

$$\frac{D}{V} = e^{2.5} \simeq 12.$$

Now D/V is the initial concentration, so for this experiment the initial concentration was 12 mg/l . For a known experimental dose $D \text{ mg}$, we can calculate that the apparent volume of distribution in this case was V litres, where $V = \frac{1}{12}D$. However, this relation would normally be used in the form $D = 12V$ to estimate the required dose for a patient for whom V is known, when an initial drug concentration of 12 mg/l (in the middle of the therapeutic range) is desired. It is found in practice that there is a simple empirical relationship between the apparent volume of distribution V (in litres) and the patient's weight W (in kilograms), namely

$$V = \frac{1}{2}W.$$

This enables one to estimate the correct dose directly from the patient's weight.

For example, for a patient weighing 50 kg the apparent volume of distribution is $\frac{1}{2} \times 50 = 25$ litres, so in order to achieve an initial drug concentration of 12 mg/l we would need to inject the patient with $12 \times 25 = 300 \text{ mg}$ of the drug.

Our model says that after the injection the concentration c is given by

$$c = 12e^{-0.17t},$$

so that it decreases steadily. After about five hours the concentration drops below 5 mg/l (the bottom of the therapeutic range). This could be prevented by administering a further injection before this time to send the concentration back up to 12 mg/l .

Such a treatment is feasible but requires many injections (one every four to five hours). There is an alternative strategy of giving the patient a slow intravenous infusion after the initial injection. This continuously replenishes the supply of drug in the blood. The mathematical model has to be changed if continuous infusion of theophylline is considered. In particular, assumption (d) must be revised, which leads to a new expression for the input-output principle and an altered differential equation.

We continue to use the value 0.17 for λ . Experiments show little variation in the value of λ from patient to patient.

Exercise 1

- (i) To make sure that you have understood what purpose is served by the modelling described in the report, describe in your own words the problem the model is intended to solve, how it arises, and why it is significant. Try to concentrate on the essential features of the problem, and in describing its significance say whether you think it is specific to the situation described in the report, or may have more general application.

- (ii) The problem specified in your answer to the first part of this question cannot be solved directly, which is where the model comes in. Fill in the blanks in the following sentence, which states what is actually being modelled.

The model describes how varies with for a given patient, as a consequence of excretion via the kidneys.

Without simply repeating the conclusions of the model, say what form you would intuitively expect this variation to take, and why.

- (iii) It is assumed in the report that the term 'concentration' is self-explanatory. How would you explain, to someone who does not regard it as self-explanatory, what it means? Your answer should probably contain a reference to one of the 'assumptions' numbered (a) to (e) which are so conveniently set out near the beginning of the report.
- (iv) The mathematical formulation of the relationship you have specified in part (ii), for any particular patient, involves two variables: define them precisely, and give their symbols.
- (v) What precisely is the model which describes how these variables are related?

Leaving aside for the moment the question of how the model is derived, let us consider the conclusions that are drawn from it.

- (vi) Sketch a graph showing how the concentration of the drug varies with time, according to the model. Does this accord with your expectations from part (ii)? What evidence is provided in the report to show that the model produces realistic results?

Insofar as the proof of this particular pudding is in the eating, you should be reasonably convinced that the differential equation

$$\frac{dq}{dt} = -\lambda q$$

gives a satisfactory model of the excretion of the drug. Let us now consider how the model is actually derived. The derivation follows directly from a very explicitly stated quantitative assumption about the process of elimination of the drug: assumption (e). So everything hangs or falls on whether this assumption is a reasonable description of how the kidneys work.

- (vii) Actually, assumption (e) cannot be correct as stated if taken literally. It asserts that the amount of drug excreted in a given time interval is proportional to the duration of the interval, however long that may be. In deriving the differential equation for q this assumption is used only for (infinitesimally) short periods of time. Is the more general form of the assumption satisfied by the solution to the differential equation, $q(t) = De^{-\lambda t}$?
- (viii) Assumption (e), even when it is modified, may seem rather arbitrary. To justify it in detail would involve knowing how the kidneys actually work, and this is a complicated business as reference to any physiology textbook will convince you. However, the important point is that it is a process of filtration, in which unwanted substances, including the drug, are filtered out of the blood. No filtration process is likely to be completely effective; but it would be reasonable to assume that every molecule of the drug has the same chance of being filtered out during its passage through the kidneys. Does this picture of the working of the kidneys lead to the result used in deriving the model?
- (ix) The model has produced a formula for the variation of the concentration of the drug with time, namely $c(t) = (D/V)e^{-\lambda t}$, which is both theoretically supported (by the argument about the working of the kidneys) and verified by experiment (so far as the qualitative predictions are concerned). But can it be used, as it stands, to solve the original problem explicitly? What further data, if any, are required to apply the model in any particular case, and how are they found?
- (x) The report gives an example of the use of the formula to find when the concentration drops below the lower limit of the therapeutic range. According to the report: 'our model says that after the injection the concentration is given by $c = 12e^{-0.17t}$... after about five hours the concentration drops below 5 mg/l'. Where does the coefficient 12 come from, and how is the time of five hours calculated?
- (xi) When a patient suffering from a bad asthma attack is admitted to hospital, he or she must be put as quickly as possible onto a routine of drug therapy which will keep theophylline in the bloodstream at a concentration lying within the therapeutic range, say between 12 mg/l and 6 mg/l, to be on the safe side. Devise a system of treatment, based on the model, which will serve this purpose. What would be the best way to present your system of treatment to the nursing staff, given that they have no time for complicated calculations in what might be a situation of some urgency, and that the relevant characteristics of the patient must be taken into account?

- (xii) The calculations in the previous two parts have taken an upper limit for the concentration of 12 mg/l, while toxicity does not set in until 20 mg/l. It is, of course, important to err on the safe side, but this looks like excessive caution on the face of it. After all, if a larger initial concentration were chosen, the patient could go for longer between boosters. Can you think of any reasons why such a low upper limit should be taken?
- (xiii) Give two suggestions for further lines of investigation which (if they can be carried through successfully) should lead to improvements in this model of intermittent administration of a drug.
- (xiv) An alternative to intermittent injection of a drug is continuous intravenous infusion, following a single initial injection. This is a method of administering the drug at a slow, steady rate. The model discussed above can be adapted to cover intravenous infusion: all the assumptions listed still hold, with the exception of assumption (d), which is replaced as follows:
- (d) after the initial injection, theophylline enters the body continuously at a constant rate.
- Derive (in the form of a differential equation) a model of the administration of a drug by intravenous infusion.

[Solution on page 39]

The second exercise is based on an extract from *TM282*. Unlike the first example, this is not designed as a formal report of a piece of modelling, but is more descriptive in nature. All the same elements are present, but you may have to work a little harder to find them. The basic ideas for this model are covered in *MST204* in *Unit 4*, but in slightly different notation: recognizing the appropriate results from that unit is part of the exercise!

2.2 Skid marks

When the police are trying to reconstruct an accident, the skid marks of a car can be very informative. From the length of the skid marks of the wheels, the police can estimate the speed at which the car was travelling before the wheels locked and the car went into the skid.

See *TM282 Unit 4*,
Section 5.2.

To evaluate the road surface the police sometimes drive a similar 'test' car with similar tyres, and cause it to skid at the same place, but at a lower speed. They then compare the skid marks they produce with the original ones. They assume that the frictional forces between surfaces are not dependent upon the speed of the car, only upon the mass of the car, the condition of the road and the type of surface. The frictional force is assumed to be proportional to the mass of the vehicle, as are any accelerating or decelerating forces due to gravity if the car is going up or down a hill. The retardation of both cars is the same according to these assumptions.

So it will be necessary to use the test vehicle's results to calculate the retardation and then to use the calculated retardation to estimate the speed of the original car.

For the test car, I shall call the initial velocity u_t and the distance s_t . The final velocity is 0, so

$$\begin{aligned} u &= u_t, \\ v &= 0, \\ s &= s_t, \\ a &\text{ is to be found.} \end{aligned}$$

The equation linking u , v , s and a is

$$v^2 = u^2 + 2as, \quad (*)$$

so $0 = u_t^2 + 2as_t,$

i.e. $a = \frac{-u_t^2}{2s_t}.$

For the original car, if it skidded to rest over a distance s at the above retardation,

$$a = \frac{-u_t^2}{2s},$$

$$v = 0,$$

$$s = s,$$

u is to be found.

Using equation (*) again,

$$v^2 = u^2 + 2as,$$

$$\text{so } 0 = u^2 - \frac{2u_t^2 s}{2s_t},$$

$$\text{i.e. } u^2 = u_t^2 \frac{s}{s_t}.$$

$$\text{So } u = u_t \sqrt{\frac{s}{s_t}}.$$

So the speed of the car before the accident can be estimated. Note that the driver's original speed would be greater than this if the car had not stopped at the end of the skid.

Exercise 2

- (i) State in your own words the purpose of this model, and say when it may be useful.
- (ii) What is the role of the test car? Why could the police not just be issued with tables giving the speed in terms of the length of the skid?
- (iii) The description of modelling that I have been trying to put across in this unit is that its basic concern is with finding a relationship between variables that specify the system under consideration; in other words, to find a formula which will enable you to calculate something if you know the value of something else, or which tells you how something varies with something else. What are the appropriate 'something' and 'something else' in this model (in words)? The second sentence of the description of the model tells you!
On the basis of what you know about skids, say what you can, in the simplest terms, about the nature of this variation.
- (iv) The term 'retardation' is used several times in the description of the model, without explanation. What do you think it means? Is it a positive or a negative quantity?
- (v) There are quite a few symbols floating around in the description of the model, but according to part (iii) we are really concerned with the relationship between just two variables. It seems to me that the symbols could conveniently be grouped into three: symbols for the fundamental variables (of which there are two); symbols for the data; and symbols which are introduced to make the calculations easier. Classify the symbols according to this scheme.
- (vi) No units of measurement are given anywhere in the definitions of the symbols. This is not the practice adopted in *MST204*, and you are advised not to follow it in this course. But does it matter in this instance?
- (vii) What is the basic model which underlies the whole discussion?
- (viii) Unlike the report of the drug therapy model, the assumptions that lead to the model are not explicitly listed here, though they are described in the text. It seems to me that there are really two fundamental assumptions: one is concerned with a special property of the acceleration (or retardation) of a car which is skidding, which permits the use of the formula in the answer to the previous part of the exercise; the other says something about the relationship between the crashed car and the test car, which enables one to use the measurements of the test car to make deductions about the crash. Can you formulate two such assumptions?
- (ix) The assumptions actually made in the text support, or give details of, these two basic assumptions. State what support there is in the text for the assumption of constant acceleration, and what support there is for the assumption that the crashed car and the test car have the same acceleration.
- (x) How do Newton's second law and the properties of kinetic friction given in *Unit 15* justify assumptions made in the text (for a level road)? (You might find it helpful to draw a force diagram for the skidding car – indeed, I hope that this idea has occurred to you without my hint.)

- (xi) The possibility that the crash took place on a slope is alluded to in the description of the model, but it is not stated explicitly whether the model applies to such a situation. Does it? You should be able to answer this question by considering whether that situation is covered by the assumptions. On the same basis, decide whether the model would apply to the following situations: a crashed Rolls Royce (assuming the police cannot afford such large and expensive cars); a crash into a very strong headwind; a crash in a shower of rain, if the road dries out before the police arrive on the scene.

- (xii) The model, so far as the crashed car is concerned, might be summarized as follows.

We want to determine the initial speed u in terms of the length of the skid s (which is the distance the car travels before coming to rest). Under the assumption of constant acceleration this is given by

$$u^2 + 2as = 0 \quad \text{or} \quad u = \sqrt{-2as}.$$

(The car's acceleration, a , is negative.)

Confirm that this result agrees with your intuitive expectations as given in part (iii).

Why is this not the end of the story, and where does the test car come in?

- (xiii) In the application of the constant acceleration formula to the original car, the equation $s = s$ appears. This does not look very informative: what do you suppose it means?
- (xiv) By combining the formulae for the crashed car and the test car, the final formula

$$u = u_t \sqrt{\frac{s}{s_t}}$$

is obtained. Says the text: 'so the speed of the car before the accident can be estimated'. If the actual skid marks are four times as long as those of the test car, and the test car was going at 40 mph when it skidded, was the crashed car breaking a 70 mph speed limit?

- (xv) Suppose that you are a police instructor: explain to a policeman about to go out on an accident investigation for the first time how to estimate the speed of the car before the accident.
- (xvi) Explain the final sentence: 'the driver's original speed would be greater than this if his car had not stopped at the end of the skid'.

[Solution on page 42]

Summary of Section 2

Reading and interpreting descriptions of models can be hard work: there is usually a lot of information to absorb, and it can be difficult to focus on what is really important. For example, it seems to me that in the drug therapy model, the name of the drug does not help one's understanding of the mathematics (though it is of crucial importance to the clinician, of course). I often think of the process of coming to terms with a model as one of excavating: digging away until you find what lies at the root of it all. In the drug therapy model it is the differential equation $\dot{q} = -\lambda q$ which lies at the root, whereas in the skid marks model it is the rule relating initial and final velocity with distance, for motion with constant acceleration. The great skill in modelling (as you will see when you come to construct models for yourself) is to formulate the fundamental equation or relationship which describes the process in which you are interested. There are several pointers in these two reports to how this is done. First, there is the importance of picking out appropriate variables and defining them carefully. Secondly, it is necessary to simplify matters in order to make progress: for example, recognizing in the drug therapy model that the mixing problem could be postponed, if not ignored. Thirdly, it is good practice to record the assumptions one makes in deriving the model: it certainly helps the reader, but more importantly, it clarifies things for you and it suggests ways of developing the model later. Fourthly, note the importance of the collection of relevant data, both to check the predictions of the model and to furnish the values of any constants that are needed to apply the model to any particular situation.

3 The skills of modelling

One thing should be clear to you by now, if it wasn't already clear before you started: mathematical modelling involves many different skills. To be good at mathematical modelling you have to be able to do some mathematics, of course; but you have to be able to do many other things as well. If you tend to think of mathematics as a set of procedures (such as the procedures for solving differential equations) then you may not think of these additional skills as being relevant to mathematics. However, most practising mathematicians, concerned with pure mathematics as well as applied, would regard that interpretation of mathematics as a travesty of the truth. One of the most powerful motivations for studying mathematics is the desire to solve *new* problems, in other words, problems for which there is no known solution procedure. The skills required in mathematical modelling are instances of general problem-solving skills. To be able to solve mathematical modelling problems is more useful, and more difficult, than being able to solve first-order differential equations by the integrating factor method. It calls on skills of invention, analysis and interpretation which apply to all sorts of problems, not just mathematical ones (pure or applied).

Here is a list of skills that may be required in the solution of a modelling problem.

- Interpreting the question.
- Choosing appropriate variables.
- Simplifying the problem and identifying its most important features.
- Specifying relationships.
- Interpreting and testing the results.
- Improving the model.

When one is tackling a modelling problem in earnest, one has to call on these skills repeatedly, in a complicated and interactive way. In particular, 'improving the model' will usually involve repeating the whole modelling process, and therefore using all the other skills over again. It seems wise to practise them individually first; this is the aim of this section. Each subsection deals with one of the skills from the list. The individual skills are illustrated with examples taken from the models in the last section. These illustrations are followed by exercises for you to do, most of which are based on models developed in previous units of this course.

3.1 Interpreting the question

In mathematical modelling the problems one faces are rarely posed in a way which can be translated directly into mathematical form. For example, from the description of the treatment of asthma in the report of the drug therapy model, one might have thought that in order to construct a model one would have to know quite a lot about drugs. However, we were able to produce a straightforward statement of the underlying problem which could serve as the basis for a model, namely

to find how the quantity of a drug in the body varies with time.

It is important to bear in mind the purpose of the model. The purpose of the drug therapy model is to devise a timetable for the administration of the drug which will keep its concentration between the therapeutic limits with the least disturbance of the patient.

This is typical of the approach that is required in order to get started on a modelling problem.

Exercise 1

The effects of double glazing were investigated in *Unit 12*. Figure 1 shows the temperature variation through a double-glazed window.

Based on this diagram, and the methods of heat transfer which produce such a temperature variation, a formula for the rate of loss of heat through the window can be derived: it takes the form

$$q = AU(\theta_{\text{in}} - \theta_{\text{out}}),$$

where q is the rate of heat loss in watts, A is the area of the window in m^2 , U is a constant called the U value of the window, with units $\text{W m}^{-2} \text{ } ^\circ\text{C}^{-1}$, and θ_{in} and θ_{out} are the temperatures of the air inside and outside the building, in $^\circ\text{C}$.

Summarize the problem which is investigated in the double glazing model, and suggest why one might want to carry out such an investigation.

Unit 12, Subsection 4.3

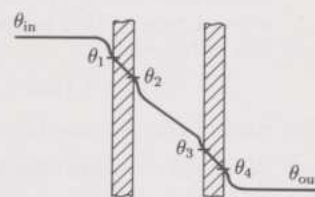


Figure 1

Exercise 2

In *Unit 15* we considered the athletics field event of putting the shot. By analysing the motion of a shot, considered as a particle moving in a vertical plane under the action of gravity alone, the range R of the shot is found to satisfy the equation

Unit 15, Section 4

$$-\frac{g}{2v_0^2 \cos^2 \theta} R^2 + (\tan \theta)R + h = 0,$$

where v_0 is the speed with which the shot is launched, θ is the angle at which it is launched, measured from the horizontal, and h is the height of the point from which it is launched above the ground.

What contribution could mathematical modelling make to the improvement of a shot-putter's performance, and how?

Exercise 3

Many shops do good trade in selling sandwiches for office workers' lunches. But no one likes to buy stale sandwiches, and quite a few fillings become unsafe to eat in a short time. The manager of a sandwich shop clearly has a problem when it comes to knowing how many sandwiches to provide for sale each day. Can you specify the problem in a form which is sufficiently precise to serve as the starting point of a mathematical model? Can you think of any other commodities, besides foodstuffs, whose sale leads to a similar problem?

[Solutions on page 44]

3.2 Choosing appropriate variables

As has already been said, the identification of the main variables is of paramount importance. If you can summarize the problem in terms of determining how one quantity varies with another, or several others, you should have no difficulty in identifying the variables. In the case of the drug therapy model there are two main variables: the first is a measure of the quantity of drug (either the mass of drug in the body, or the concentration of drug in the bloodstream); the other is the time since the drug was administered. In the skids model the problem is to find how the length of the skid depends on the speed of the car when skidding began, from which the variables are obvious. It pays to be careful in defining variables ('time since the drug was administered', not just 'time'). One should give units, even though I haven't done so here.

Exercise 4

In *Unit 4* we posed the following problem.

Unit 4, Subsection 3.3

'An object which is initially at rest is dropped from the Clifton Suspension Bridge and falls into the River Avon, 77.0 metres below. Assuming that the only force acting on the object is the force of gravity, find the time of fall.'

Select two main variables for this problem, and define them carefully.

Exercise 5

The modelling of a mortgage was discussed in the first TV programme of the course. The programme dealt with the repayment of a loan of £20 000 over a period of 25 years at an interest rate of 15% per annum. A recurrence relation for u_r , the amount owed after r years, was obtained:

$$u_{r+1} = 1.15u_r - 12M,$$

where M is the monthly repayment.

What are the main variables in the mortgage problem?

Unit 1, Subsection 4.1

Exercise 6

New supermarkets open so frequently that I can't help wondering how the managements of supermarket chains decide where to site new stores. One fundamental factor in the choice of site must be the catchment area. Can you define variables which would be useful in setting up a mathematical model which could be used to decide the position of new supermarkets?

[Solutions on page 44]

3.3 Simplifying the problem and identifying its most important features

The skid model depended on the assumptions that the acceleration of a skidding car is constant, and that for given conditions different cars have the same acceleration while skidding. These assumptions follow from a well-established theory of sliding friction, but hold only if (for example) air resistance is ignored. To ignore air resistance is justified on two counts: first, its effects are probably small compared with those of sliding friction; secondly, the resulting model is relatively easy to analyse. In modelling one should always look for as simple a model as possible consistent with the salient features of the problem (to have ignored the effects of friction completely would obviously have been counter-productive). But it is important to be clear about the simplifying assumptions that have been made in order to arrive at that model. This is necessary to ensure that anyone who may be interested can follow how the model has been derived. Furthermore, if you have an explicit list of the assumptions you have made, and you wish to improve your model, you have an obvious place to start: review the assumptions and ask which can and should be modified or relaxed.

Exercise 7

What simplifying assumptions would you make in setting up the equation of motion in the Clifton Suspension Bridge problem described in Exercise 4?

Exercise 8

What assumptions must be made to derive the equation $q = AU(\theta_{\text{in}} - \theta_{\text{out}})$ in the double glazing model of Exercise 1?

Exercise 9

How much of a car's windscreen or rear window is actually cleaned by the windscreen wiper or wipers depends on the length of the wiper blade(s) and the positioning of the wiper(s) on the windscreen. It is obviously desirable that the area cleaned should be as large as possible. What simplifications would you make in order to set up a model of the operation of windscreen wiper(s) that would help to determine the best arrangement of the wiper(s)?

[Solutions on page 45]

3.4 Specifying relationships

The use of the input-output principle to derive the fundamental differential equation in the drug therapy model is a good illustration of quite a common technique; we used the same technique in Unit 3 when developing models for populations. Of course, most models of mechanical systems involve the use of Newton's laws, though the use may be hidden, as in the skid model.

Exercise 10

Explain, in your own words, how the recurrence relation in the mortgage model (described in Exercise 5) is derived. In the problem considered in *Unit 1*, specific values for the interest rate and the loan are given. But all repayment mortgages work in the same way, so it should be possible to write down a recurrence relation for any such mortgage. Can you do so? You will have to introduce into the model some parameters which you will have to define.

Exercise 11

In *Unit 8* a highly simplified model of the suspension of a car was discussed. The basic features are shown in Figure 2. The car body is represented by a particle, which is supported by a perfect spring and a damper, together representing the suspension. To test the suspension, a potential buyer of the car depresses the body a certain distance below its normal position (with the car stationary) and releases it.

Describe how the equation which determines the subsequent motion of the car body is obtained.

Exercise 12

A supermarket where I frequently go shopping is surrounded by car parks as shown in the plan below.

Unit 8, Subsection 1.4,
Exercise 7

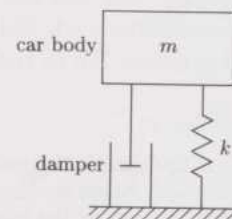


Figure 2

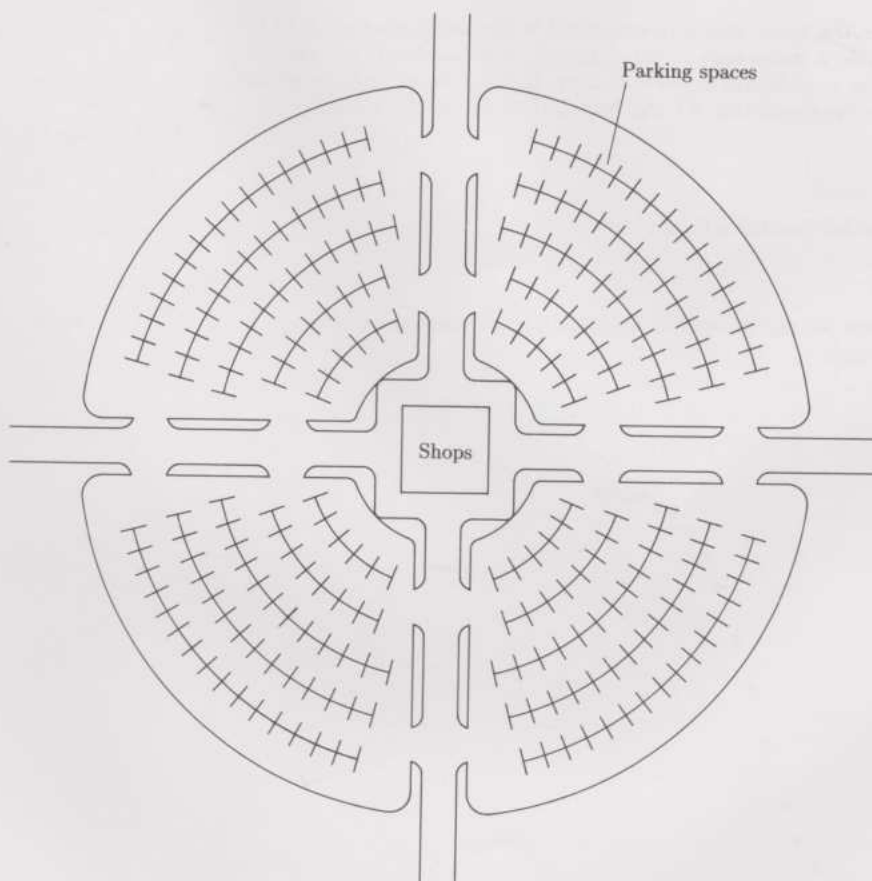


Figure 3

When I go there I can never decide where to park. The problem arises because the closer to the centre one goes, the longer one has to spend driving around looking for an empty parking space; if one parks far away from the centre, however, one has to spend a long time walking to the shops. It seems to me that I could try to resolve this difficulty as follows. I shall try to find how the time between arriving at the car park and walking into the shopping area depends on where I park, and then work out how far away to park to minimize this time. Let t be the time between arriving at the car park and walking into the shopping area, and x the distance I park from the shops. The time it takes to find somewhere to park is a decreasing function of x : I suspect it is inversely proportional to x .

Formulate a relationship between t and x that could be used to solve the problem.

[Solutions on page 45]

3.5 Interpreting and testing the results

Here there are several points to be emphasized. It is usually worth checking that the model predicts the kind of results that one would expect from common sense and from experience. If possible, one should validate the model by comparing its predictions with data from an experiment or other reliable source. It is good practice to try to reformulate the results so that this check turns into something simple like drawing a straight line, as in the drug therapy model. The model may not directly provide the information you want to solve the initial problem, so you may have to do some further mathematics (solve a differential equation, or find a maximum or minimum, for example). But you may also require some data to give an explicit numerical solution to the problem – such as the value of λ in the drug therapy problem. The beauty of the skid problem lies in the use of the test car to provide these extra data.

Exercise 13

Unit 3 dealt with the growth of populations. An initial attempt to model the growth of a population, based on the principle that the change in the population in a short interval of time is proportional to the population size at that time, leads to the equation

$$\frac{dP}{dt} = \lambda P,$$

where P is the population size, t is the time, and λ (a constant) is the proportionate rate of increase of the population. To obtain a more sophisticated model, it is assumed that the proportionate rate of increase of the population is not constant, but decreases linearly as the population size increases. The resulting equation for the population size is the logistic equation,

$$\frac{dP}{dt} = aP \left(1 - \frac{P}{M}\right)$$

(where a and M are constants), which has the solution

$$P = \frac{M}{1 + (M/P_0 - 1)e^{-at}},$$

with P_0 the initial population. Data for the growth of a population of yeast cells, when plotted, produced the following graph.

Unit 3, Section 2

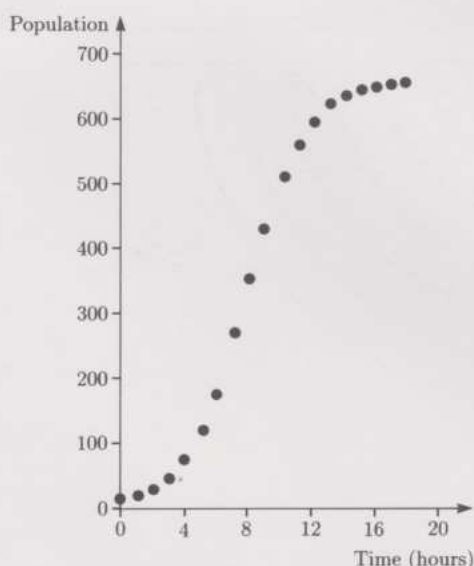


Figure 4

What qualitative features of the solution to the logistic equation suggest that it is an appropriate model for population growth? How can the solution be rearranged to provide a straightforward way of checking the model against data such as those given for the population of yeast cells?

Exercise 14

The recurrence relation for the mortgage problem quoted in Exercise 5 was used to find the value of M for which the loan would be repaid exactly after 25 years; the value obtained was

$$M = 257.8324 \dots$$

It was discovered that a monthly repayment of £258 (obtained by rounding the value of M to the nearest pound) leads to an eventual overpayment of £427.

Describe exactly how the recurrence relation for the mortgage problem is used to find the monthly repayment, taking into account the practical problems of implementing the solution (taking the effect of rounding as an example) as well as the mathematical ones of deriving it.

Exercise 15

An attempt to model cooking times, which was the subject of a television programme in *M101* and *MS284*, leads to the following result. Consider a spherical body of mass m , initially at room temperature, which is placed in an oven maintained at a much higher temperature. The aim of the model is to find the time it takes for the centre of the body to reach a predetermined temperature (the temperature at which food is deemed to be cooked). The model predicts that this time t is given by

$$t = km^{2/3},$$

where k is a constant. What checks can you devise to confirm that this is a reasonable result? How would you use the result to find how long it would take to bake a potato?

[Solutions on page 46]

*M101 Block V Unit 1,
Section 1.4
MS284 Unit 14, Section 4*

3.6 Improving the model

The points that were ignored in simplifying the model can now be reconsidered. In the case of the drug therapy model it was suggested that the initial mixing of the drug with the bloodstream should be considered. A less demanding modification of the skid model was briefly raised in the final part of the exercise, which asked about the effect of relaxing the implicit assumption that the car stops at the end of the skid. Since the purpose of the exercise is to find incontrovertible evidence that the driver was speeding when the accident occurred, a qualitative answer was enough.

Exercise 16

Explain why the exponential model of population growth, described in Exercise 13, is usually inadequate. Discuss the modifications which lead to the logistic model.

Exercise 17

In the Clifton Suspension Bridge problem (Exercise 4), an initial model is proposed in which velocity-dependent forces are ignored; this model is superseded by a better one in which such forces are taken into account. What is the effect of this on the result?

Exercise 18

When a drug is taken by mouth, rather than being given by injection, there is a delay before it gets into the bloodstream because it has to pass to the stomach and be absorbed. Can you devise a model of the change in concentration over time of a drug taken by mouth?

[Solutions on page 47]

Summary of Section 3

The solution of a modelling problem will usually require the use of the following skills.

- Interpreting the question.
- Choosing appropriate variables.
- Simplifying the problem and identifying its most important features.
- Specifying relationships.
- Interpreting and testing the results.
- Improving the model.

Identifying and practising these skills at every opportunity will pay dividends when you come to do some modelling for yourself.

4 Putting it all together

In previous sections you have learnt how to interpret models and you have practised the individual skills required in modelling. You must now begin to integrate these skills, so that you can tackle modelling problems yourself. In this section I am going to talk my way through a modelling problem, describing how I would think about it; I hope that this will give you some guidance as to what is involved in doing one's own modelling. Most of this section consists of reading only, though there are some exercises for you to do at the end to help you consolidate the points that have been made; but I hope you will read in an enquiring frame of mind, and keep asking yourself at each stage what you would do next if you were tackling the problem. My general line of attack is similar to that already brought out in the discussions of the drug therapy and skidding models in Section 2; so you should probably find yourself always one step ahead of me as you read this passage.

In the next section you are asked to put the ideas into practice, by doing a piece of modelling yourself, with guidance.

The question I am going to consider is this:

- Immediately after a cup of tea is poured out, it is usually too hot to drink.
- How long does it have to be left to cool before it is drinkable?

This is not, on the face of it, a terribly important question, and falls within the category of questions one asks out of curiosity rather than because the answer is likely to be useful. However, there are many different situations in which cooling plays a significant role, some of which are of considerable importance; you should not find it difficult to think of industrial processes in which cooling plays an important part, for instance. The cooling of a cup of tea is a familiar and simple example, which involves the principles which apply in most cooling problems.

I am now going to describe the steps I would go through in trying to answer this question.

The problem is concerned with *how hot* a cup of tea is: the first thing to do is to find a precise way of stating what this means. How does one measure how hot something is? With a thermometer. In other words, the problem is concerned with the temperature of the tea. Furthermore, the question asks *how long* it takes for something to happen, so time is going to be of major importance. The guts of the problem, therefore, is to find how the temperature of the tea varies with time, and the two main variables can be identified already: they are

- θ , the temperature of the tea, in $^{\circ}\text{C}$;
- t , the time since the tea was poured out, in seconds.

My interpretation of the problem is that θ is a function of t , and I have to find out as much as I can about this function – if possible, I want to express $\theta(t)$ explicitly in terms of t . I imagine that, if I can do so, the expression will contain various constants (which I tend to call ‘parameters’) which represent various relevant physical characteristics of the system: the initial temperature of the tea and the temperature of the surroundings are obvious examples. More will no doubt present themselves as my ideas develop.

Before I try to analyse the cooling process in detail, I shall ask myself how, intuitively, I would expect the temperature to vary with time. The first and most obvious point is that anything which is hotter than its surroundings cools down: so the temperature of the tea will progressively decrease as time passes. How can I express this mathematically? By saying that $\theta(t)$ will be a decreasing function. But can the temperature keep on decreasing indefinitely? It is surely a matter of experience that however long one leaves it, one's tea never gets colder than its surroundings; but if one waits long enough, its temperature will get as near that of the surroundings as makes no difference. What would a graph of temperature against time be likely to look like? Figure 1 shows a couple of possibilities.

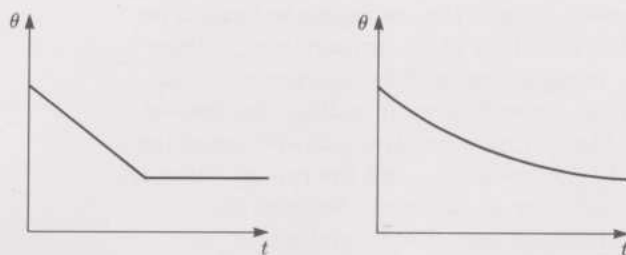


Figure 1

Which is the more likely? The second is more likely than the first, surely: physical processes tend to be smooth.

It occurs to me that I could make an educated guess at the function represented by the second graph. Its shape is that of a decaying exponential, but it is moved up a little, parallel to the vertical axis. Can I use this to guess a formula for $\theta(t)$? It suggests that the relation between θ and t takes the form

$$\theta(t) = a + be^{-\lambda t},$$

where a , b and λ are constants. Here $be^{-\lambda t}$ is the basic exponential: to obtain the shape shown in the figure, b and λ must both be positive. The addition of the constant a corresponds to moving the exponential up parallel to the vertical axis. This formula for $\theta(t)$ gives a partial answer to the question 'how does θ depend on t ?'; it would be of more use, however, if I could interpret a , b and λ in terms of the parameters of the problem. This can easily be done for a and b by thinking about initial and (if you see what I mean) final conditions. First, putting $t = 0$, I obtain

$$\theta(0) = a + b.$$

But $\theta(0)$ is the initial value of the temperature of the tea, its temperature when it is poured out; it would fit in with the usual practice if I were to denote this by θ_0 . On the other hand, I suggested that in the long term (as $t \rightarrow \infty$) the temperature of the tea will tend to that of its surroundings, say θ_s ; but as $t \rightarrow \infty$, $e^{-\lambda t} \rightarrow 0$, so this tells me that $a = \theta_s$. I can now see that $b = \theta_0 - \theta_s$, and so

$$\theta(t) = \theta_s + (\theta_0 - \theta_s)e^{-\lambda t}.$$

So far so good, but λ is still undetermined. What is more, the formula was obtained by a piece of heroic guesswork. How could I confirm the accuracy of my guess, and how can I find the value of λ ? One way would be by experiment: take a cup of tea and measure its temperature periodically. If I were to plot the results (measured temperature against time) I would hope to find that the points lay on a graph like that in the previous figure. How could I check the model more effectively? By making a plot of $\log_e(\theta(t) - \theta_s)$ against t , because the model predicts that the result would be a straight line, of slope $-\lambda$. Not only would this give confirmation of the model, but it would provide the value of λ . I expect you will have noticed the similarities between this and the drug therapy model.

This procedure has its limitations, however. It would confirm the result for a particular cup of tea, and might reasonably be supposed to provide supporting evidence of the correctness of the model for cups of tea in general. But because it is basically empirical, it is of rather limited validity. My aim must surely be to find a method of predicting the time of cooling for any of the many situations in which the knowledge might be useful, and having to carry out a new set of measurements in each case cannot really be described as a satisfactory approach.

It would be better to find an analytical derivation if possible. To do so I must think more carefully about the process of cooling. So at this point I must make a mental effort and try to recall what there is in *Unit 12* that might be relevant and useful. That unit is about the flow of heat, and I haven't thought about heat at all yet in my consideration of the problem. The first question is 'how are heat and temperature related?'. The essential points seem to be these: where there is a difference of temperature, heat will flow (from the warmer to the cooler place); and when something loses heat, and the loss is not made good, then its temperature will fall. So the

following picture of the processes underlying the cooling of the tea begins to form in my mind: since the temperature of the tea is higher than that of its surroundings, it loses heat continuously; since the tea is losing heat, its temperature falls continuously. The greater the difference in temperature between the tea and its surroundings, the greater the rate at which heat is lost, and the greater the rate at which the temperature of the tea falls. Conversely, as the tea cools, it loses heat more slowly, and the rate at which it cools slows down. In the long term, there is practically no difference between the temperature of the tea and that of its surroundings, no loss of heat, and no further cooling. It sounds as though this mechanism will reproduce the qualitative features of the cooling process that have been discussed already, which is promising. So next I shall try to turn these ideas into a mathematical model.

Three ways in which heat flows from one place to another are discussed in *Unit 12*, as I am sure you will remember: they are radiation, conduction and convection. Which of these will apply in our case? Radiation is connected, in my mind at least, with things like electric fires and the Sun; if I put my hand near a cup of tea, I won't feel any warmth worth speaking of unless I actually touch the cup or the tea, so there can't be much heat being radiated and it seems safe to ignore it. Both of the other methods of heat transfer will apply, though: heat is conducted through the cup, and is convected directly from the surface of the tea and also from the outer surface of the cup. To keep things simple it would be a good idea to concentrate on just one of these forms of heat loss (conduction or convection), if possible: this will be justifiable if it seems likely that heat is lost much more rapidly by one route than the other. The effect of ignoring the less significant of the two will be that the answer may be a little inaccurate, but it is better to develop a simple model which you can be confident will actually produce an answer, than to develop one that is so complicated that you cannot use it. Improvements can be made bit by bit, but only if you have something which actually works to improve on.

The outside of a cup containing hot tea is a lot cooler than the tea itself: after all, one reason why china, for example, is used to make cups is that it is a good insulator. So heat is probably lost much more rapidly by convection from the surface of the tea than it is by conduction through the cup. To begin with, therefore, I shall ignore all forms of heat loss except convection from the surface of the tea. Now, what do I need to know about the tea and its surroundings, in addition to the difference in their temperatures, in order to calculate the rate of loss of heat by convection? I need to know the area of the surface of the tea: let me call this $A \text{ m}^2$. I also need to know the convective heat transfer coefficient: let me call this $h \text{ W m}^{-2} \text{ }^\circ\text{C}^{-1}$. Then if q is the rate of loss of heat (in watts), and the temperature difference between the tea and the surrounding air is $\theta - \theta_s$,

$$q = hA(\theta - \theta_s).$$

That has made a start on the first half of the cooling process: I have a model of the way the temperature difference between the tea and its surroundings leads to a flow of heat from the tea to the surroundings. The other half of the process is concerned with the resulting fall in the temperature of the tea. Near the beginning of *Unit 12* there is a formula which relates a change in the internal energy of a body to the consequent change in its temperature. You may remember this formula, which is given in the form

$$E_2 - E_1 = mc(\theta_2 - \theta_1).$$

In *Unit 12* this formula is discussed in the context of the heating of water in a kettle: the heat transferred to the water by the kettle raises its internal energy from E_1 to E_2 , with resulting increase in temperature from θ_1 to θ_2 ; m is the mass of the water, and c a constant, characteristic of the substance being heated, called its specific heat. This formula applies equally to the fall in temperature resulting from a loss of energy. In our case, it can be rewritten as follows: when a mass $m \text{ kg}$ of tea loses δE joules of heat energy, its temperature falls by $\delta E/(mc) \text{ }^\circ\text{C}$, where c is the specific heat of tea, in $\text{J kg}^{-1} \text{ }^\circ\text{C}^{-1}$.

I now have to put the two parts of the process together. There is, however, a problem: the formula for the rate of loss of heat by convection applies only in the steady state, that is, when the temperatures concerned remain constant in time; this is definitely not

Strictly speaking, this formula holds only in the steady state. We shall deal with this point shortly.

Unit 12, Subsection 1.3

true of the situation under consideration – indeed, the whole point of the exercise is to find how the temperature of the tea *varies* with time. But I can get round this by a standard device, which you have met once already in this unit. I shall consider what happens in a short interval of time of length δt , during which the temperature of the tea remains effectively constant. I can then use the steady-state result as an approximation. It is an approximation which improves as $\delta t \rightarrow 0$. The price I have to pay is that I shall finish up with a differential equation.

In the interval of time from t to $t + \delta t$, the *rate* of loss of heat by convection, using the steady-state approximation, is

$$q(t) = hA(\theta(t) - \theta_s).$$

The quantity of heat lost is therefore

$$\delta E = q(t)\delta t = hA(\theta(t) - \theta_s)\delta t.$$

When the tea loses this amount of heat energy its temperature drops by $\delta E/(mc)$, and so

$$\theta(t + \delta t) = \theta(t) - \frac{\delta E}{mc} = \theta(t) - \frac{hA}{mc}(\theta(t) - \theta_s)\delta t.$$

It follows that

$$\frac{\theta(t + \delta t) - \theta(t)}{\delta t} = -\frac{hA}{mc}(\theta(t) - \theta_s).$$

Now let $\delta t \rightarrow 0$ to obtain

$$\frac{d\theta}{dt} = -\frac{hA}{mc}(\theta - \theta_s).$$

This is my analytically derived model of the cooling process.

Once you have a differential equation, there is one obvious thing to do: try to solve it. But before doing so, one might ask whether it predicts the right kind of behaviour for θ . We expect θ to decrease with t so long as $\theta > \theta_s$, and to decrease more slowly the nearer θ is to θ_s . Is the differential equation consistent with this behaviour? Well, so long as $\theta > \theta_s$, we have $d\theta/dt < 0$, so θ is decreasing; and the closer that θ is to θ_s , the smaller is $|d\theta/dt|$. Note that $\theta = \theta_s$ is a constant solution, or equilibrium state (like the terminal velocity for a particle falling under air resistance, or the equilibrium population size for a population growing according to the logistic equation). So the equation does appear to predict the right kind of behaviour.

How is it solved? One way is to use the integrating factor method (the integrating factor is $e^{\lambda t}$, where $\lambda = hA/(mc)$). Another is to use separation of variables (divide through by $\theta - \theta_s$). Here is a third method, quicker than either of the others. Put $\phi = \theta - \theta_s$: then $d\phi/dt = d\theta/dt$ (because θ_s is constant), so the equation becomes

$$\frac{d\phi}{dt} = -\frac{hA}{mc}\phi = -\lambda\phi, \quad \text{where } \lambda = \frac{hA}{mc},$$

an equation whose solution you should know (if only because we have met it before in this unit, in the drug therapy model):

$$\phi(t) = \phi_0 e^{-\lambda t}.$$

In terms of the original variable θ we have

$$\theta(t) - \theta_s = (\theta_0 - \theta_s)e^{-\lambda t}, \quad \lambda = \frac{hA}{mc}.$$

I should just like to be sure this makes sense. The formula agrees with the one I guessed before by thinking about the graph of θ against t . The temperature θ tends to θ_s in the long term. The temperature will fall more quickly if the surface area A is increased: that seems right, because in that case there will be more room for the heat to escape. On the other hand, if you have more tea it will cool more slowly: a greater mass will hold more heat, and the loss of a given amount of heat will result in a smaller drop in temperature. So the answer does appear to make sense.

It remains to answer the original question: how long will it be before the tea is drinkable? How do we set about answering this? So far as purely mathematical

These ‘equations’ are, strictly speaking, only approximations; however, they get more and more nearly true as δt gets smaller.

If I had anticipated this outcome, I would have used ϕ as my main variable in the first place!

processes are concerned, it is simply a matter of setting $\theta(t)$ equal to the temperature at which tea becomes drinkable, say θ_d , and solving for t : this gives

$$\theta_d - \theta_s = (\theta_0 - \theta_s)e^{-\lambda t},$$

or
$$\frac{\theta_d - \theta_s}{\theta_0 - \theta_s} = e^{-\lambda t},$$

so
$$t = \frac{1}{\lambda} \log_e \left(\frac{\theta_0 - \theta_s}{\theta_d - \theta_s} \right) = \frac{mc}{hA} \log_e \left(\frac{\theta_0 - \theta_s}{\theta_d - \theta_s} \right).$$

But this isn't the end of the story, is it: in order to obtain an explicit numerical answer, the values of the various parameters in this formula need to be known. Here a list would be useful – or, even better, a table like the following. As well as listing the meanings of the symbols and the units of measurement, I have asked myself in each case how the value is to be obtained (as no doubt you have too).

Symbol	Definition	Units	How obtained
m	Mass of tea	kg	Weigh a full and an empty cup – use cold tea to avoid accidents!
c	Specific heat of tea	$\text{J kg}^{-1} \text{ }^\circ\text{C}^{-1}$	Use value for water given in <i>Unit 12</i>
h	Convective heat transfer coefficient between tea (water) and air	$\text{W m}^{-2} \text{ }^\circ\text{C}^{-1}$	Look up in book of physical constants
A	Area of surface of tea	m^2	Measure diameter of top of cup
θ_0	Initial temperature of tea	$^\circ\text{C}$	Measure with thermometer: presumably just below boiling point
θ_s	Temperature of surrounding air	$^\circ\text{C}$	Measure with thermometer
θ_d	Temperature at which tea becomes drinkable	$^\circ\text{C}$	Sample, then measure with thermometer

Once the values of these parameters have been found, the rest is arithmetic – or is it? Here are some real-world values (actually for a mug rather than a cup).

Quantity	Value
m	0.25 kg
c	$4200 \text{ J kg}^{-1} \text{ }^\circ\text{C}^{-1}$
h	$250 \text{ W m}^{-2} \text{ }^\circ\text{C}^{-1}$
A	0.005 m^2
θ_0	$80 \text{ }^\circ\text{C}$
θ_s	$20 \text{ }^\circ\text{C}$
θ_d	$60 \text{ }^\circ\text{C}$

With these values my calculator gives the following answer:

$$t = 340.5906908.$$

It is very tempting to conclude that the time taken for the tea to cool enough to be drinkable is 340.5906908 seconds, or 5.676511513 minutes. What would be wrong in doing so? I have derived a model, based on good physical laws and principles; I have correctly calculated the result this model predicts, using reliable data; so the answer must clearly be right. But can one really trust an answer given to 10 significant figures – an answer that purports to have a far greater degree of accuracy than any watch on the market? The data are given to at most two significant figures, so we can scarcely expect the result to be more accurate. In any case, the model itself is approximate, as we know: the effects of conduction, for example, have been ignored. It would be much more honest to state the result as ‘the time taken for the tea to cool enough to be drinkable is about 6 minutes’.

The consequence of ignoring the effects of conduction should also be considered. The values of convective heat transfer coefficients are difficult to determine accurately, and they depend on the precise conditions in a way that is difficult to predict: this is implicit in the enormous ranges of values quoted in Table 3 on page 17 of *Unit 12*. Whether or not the tea is stirred, and whether or not there is a draught in the room in

which you are drinking it, could seriously affect the value of h . I suspect that the correction resulting from taking account of conduction through the cup would be less than the error resulting from the indeterminacy in the value of h , so that from the practical point of view there would be little to be gained from developing the model in that direction. However, it is worth asking whether the inclusion of the effects of conduction would change the qualitative behaviour of the model, and whether it would tend to increase or decrease the estimate of the cooling time. Since the formula for the rate of heat loss by conduction takes the same form as that for the rate of heat loss by convection – in each case the rate of heat loss is proportional to the difference between the temperatures of the tea and the surrounding air – the only change to the model will be that hA will be replaced by another term. Moreover, the new term will be larger than hA (since it will incorporate additional heat loss), so the effect of the inclusion of conduction in the model will be to decrease the estimate of the cooling time slightly, as one would expect.

I do not seriously suggest that whenever one has a cup of tea one should calculate how long to leave it before it is cool enough to drink. The benefits of the foregoing exercise were intended to be entirely pedagogical. You should have learnt two things from it. One is a new (so far as this course is concerned) piece of information about heat, namely the law that governs the cooling of hot bodies. This is known as Newton's law of cooling, and can be stated briefly as follows: the rate of decrease of temperature is proportional to the temperature difference. The other is some further insight into the nature of modelling. I'd like to summarize the points that have arisen which, in my opinion, throw useful light on the nature of modelling, in the form of a collection of questions that one is likely to have to ask oneself at various stages when one is engaged on a modelling problem.

Questions to ask yourself while modelling

What process will I have to model in order to answer the question posed?

Here the question asked was 'how long does it take for a cup of tea to cool?'; the process to be modelled is the transfer of heat between an object and its surroundings, and the consequent changes in the object's temperature. I put it in this rather vague way so that you can see that the model is actually much more general than the particular situation apparently under investigation: the same principles should apply to the warming of cold objects as to the cooling of hot ones, for example.

What are the key variables I shall have to use in my model?

'How does one measure how hot something is?' turned out to be a useful pointer to temperature, one of the two key variables in this case. It is very easy to overlook time as a variable, simply because it is so obvious: it is almost a tautology to say that time is likely to be a key variable in a problem concerned with how long it takes for something to happen. Likewise, spatial variables (coordinates, for example) are bound to be required in problems concerned with spatial variation or change in position. In most of the modelling problems you have met or will meet in this course, the key variables will be few in number – usually two or three. (The accident frequency model with which this unit started is untypical in this respect – but it was chosen to make a point.)

Can I formulate the underlying modelling problem in terms of how one of my variables depends on the other(s)?

If you have a clear idea of the key variables you will probably have already answered this question, and vice versa; the only extra point is to decide which is the dependent and which the independent variable, as we say in the trade. My succinct formulation of the cooling problem is 'how does the temperature of the tea vary with time?'; here temperature is the dependent variable, time the independent one. Forms of relationship other than functional dependence arise in modelling problems, of course, but the question, suitably modified, is still relevant.

What do common sense and experience tell me about the qualitative aspects of the relationship between my key variables?

I got quite a long way with the cooling problem by this form of guesswork. You won't always be so lucky, but thinking about the problem in this manner may suggest, for example, a hypothesis to test empirically, which will be some sort of progress even if you can't see how to develop a theoretical model. In many cases (such as the accident frequency problem) this will be the only way of getting anywhere. When you can develop a theoretical model, on the other hand, observations along the lines of ' y must be an increasing function of x ' provide useful qualitative checks. (But always remember that neither common sense nor experience is completely reliable; indeed, most of the really important advances in science have been due to someone showing that what everyone else thought was common sense actually isn't.)

What would I learn from an experiment?

If you are contemplating some kind of empirical approach, you must be clear what you are looking for. Experiment, or the collection of masses of data, for its own sake, is not good modelling. There are usually two reasons for experimenting, or collecting data by some other means: one is to test a hypothesis about the nature of a relationship between variables; the other is to find the values of parameters which occur in the model. In the cooling problem an experiment could have been used to confirm that the temperature falls exponentially with time, and also to find the value of the parameter λ which measures the rate of cooling. These are examples of the two different possible uses.

How can I find a relationship between my variables theoretically?

There is no simple answer to this question, no way of producing a Procedure Box for modelling. But there are several tips and suggestions that I can give.

- First, it can often help if you have some idea from the start what kind of model you are expecting to obtain. An ordinary algebraic equation? An equation involving trigonometrical functions, or exponentials? A differential equation? If so, what will its order be?
- Look for relevant laws or principles, such as the input-output principle, the law governing heat transfer by convection, Newton's laws, and so on. The idea of considering 'what happens in a short period of time of length δt and letting δt tend to zero' should be mentioned here too.
- Ask yourself whether you have previously met other modelling problems with similar features (the drug therapy problem has surprisingly many similarities with the cooling problem, for example).
- Try to develop a simple picture of the processes involved, in your own terms and language. This could take the form of a metaphor (the kidneys are like a filter). A diagram is often useful – essential, I would say, in the case of mechanics problems, to show the forces.
- Simplify like mad. In most real-life situations there are many processes going on simultaneously. Some will be quite irrelevant to the problem in question (whether the beverage is tea or coffee); others can be ignored, at least initially, because their effects are relatively insignificant (radiation and conduction compared with convection). Your aim should be to identify just the one or two most important mechanisms at work, and model those. If this leads to a somewhat superficial and inaccurate model, this doesn't matter very much, since it is always possible to try and incorporate additional features later. It is much better to have a simple model which you can subsequently try to improve, than to be prevented from making a model at all because you are trying to cover everything.
- Lists can be helpful, such as the list of the different mechanisms for heat transfer mentioned in the solution to the cooling problem. You can work through the items

on the list and decide which have the most significant effects. If you can put them in order of importance, so much the better (convection first; conduction second; radiation nowhere). I have to admit that this order would be more convincing if I had done some rough calculations to estimate the relative sizes of the rates of loss of heat by each of the three methods. But, assuming that I'm right, a clear policy emerges: model the heat loss by convection first, incorporate conduction if necessary next, but don't bother with radiation at all. There is another benefit to be gained from drawing up such a list: when reporting on your modelling to someone else it is very helpful to the audience to be explicit about which mechanisms you have taken account of, which you have ignored, and why – in other words, to make explicit your assumptions in the manner of the report on the drug therapy model – and a list will help you to do this.

- Don't be afraid to introduce new variables, such as q in the cooling problem, and parameters such as h and A . (The distinction here is between the variables θ and t , which would appear in any cooling problem, and the parameters, h , A , m , c , θ_s , which are constants for any particular cooling problem, but whose values vary between different cooling problems. The rate of loss of heat, q , and the energy change, δE , are variables, but have a different status from θ and t : q and δE are introduced to make the derivation easier, but they are hidden, or subsidiary, in the sense that they don't appear in the final model.) As you think about the processes involved, the need for these new variables and parameters should become apparent. Don't let them multiply unchecked, however: if you are running out of letters of the alphabet, you probably haven't simplified enough. Do keep a list of symbols you introduce, with their definitions and units: there is nothing more irritating than deriving a beautiful, complicated formula and then realizing you have forgotten what α_ω means.

When I have derived a model, what do I do next?

Now is the moment to remind yourself what the problem is that you are trying to solve, and to reformulate the problem in terms of the mathematical model you have obtained.

How do I use my model to solve the original problem?

Do some mathematics! If you have obtained a differential equation, or a recurrence relation, solve it. If your problem asks you to find a maximum or minimum, differentiate and set the derivative to zero. Of course, what mathematics you do will be determined by the problem that was originally posed. Remember that numerical answers are usually required in the real world, and so this part of the exercise involves thinking about how one finds the values of any parameters that are needed to obtain a numerical answer. Remember also to give the answer to a level of accuracy which is appropriate to the problem, the model and the data, to the best of your judgement; in particular, don't give the appearance of claiming greater accuracy than the model and the data can support.

How can I be confident I'm on the right track?

This is a good point at which to make some qualitative checks. Confirm that your model predicts the kind of qualitative behaviour you expect, including, for example, what happens in the long term and in limiting cases (as I did in the air resistance problem and the logistic model). Try the effect of varying the value of a parameter, and seeing if the answer makes sense (as I did for m and A on page 31 in the cooling problem).

How can I check that the model corresponds to reality?

You might consider conducting an experiment, as described in the drug therapy model for example; or you might look for confirmatory data from elsewhere, such as reports of other investigations of a similar problem.

How can I improve my model?

Here is where your list of assumptions will be useful: look at what you decided to leave out the first time round, and consider how it can be incorporated in the model.

In *Unit 3* we used the following diagram to summarize the various activities involved in modelling.

M101 Block V Unit 1,
Section 1.4
MS284 Unit 14,
Subsection 4.1

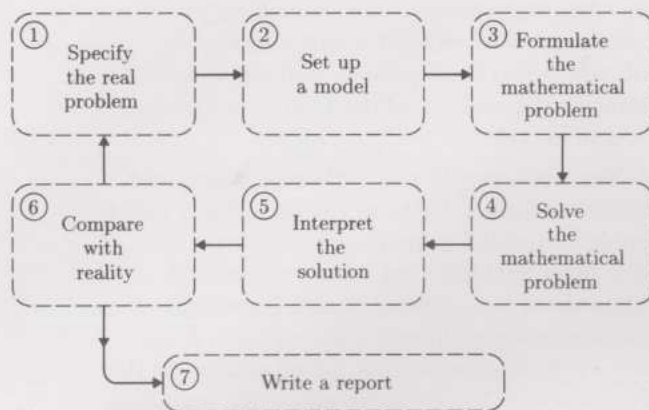


Figure 2

The list of questions and answers above constitutes an attempt to describe in greater detail what may be involved in the first six of these seven activities. (We shall discuss writing reports in the *Guide to Mathematical Modelling*.) The diagram can help in modelling, but is intended only to remind you of the things that you may have to do; as we pointed out in *Unit 3*, it is not intended to be a straightjacket.

Here are a couple of problems on the theme of heat transfer, for you to think your way through to be sure you have understood the points which have been raised in this section. You are not being asked to solve the problem in either case, just to think how you would set about solving it. (For the sake of completeness, the solutions to these exercises contain more detail than you are expected to provide in answering them.)

Exercise 1

I am fond of fresh milk, provided that it's cool: I do not gain much pleasure from drinking milk near room temperature. I keep my milk in the refrigerator, of course; I want to know for how long a bottle of milk remains pleasantly drinkable after it has been taken from the refrigerator. How would you set about solving this problem?

Exercise 2

A thermostatically controlled domestic central heating system works roughly as follows. When the temperature in the house falls to a predetermined lower value, the heating is turned on, and delivers heat at a constant rate, until the temperature rises above a predetermined higher value. The heating then goes off and the house cools, until the cycle begins again. How would you set about modelling this situation?

[Solutions on page 49]

Summary of Section 4

Here is a list of points it may be helpful to consider when trying to develop a mathematical model. It is based on the questions that I have suggested you ask yourself while modelling, but I have grouped them so that they correspond to the stages of the modelling diagram.

1. Specify the real problem
 - Decide what process you will have to model in order to answer the question posed.
2. Set up a model
 - Pick out the key variables that will have to be used in the model.
 - Try to formulate the underlying modelling problem as 'how does one of my variables depend on the other(s)?'.
 - Use your common sense and experience to say what you can about the qualitative aspects of the relationship between your key variables.
 - Formulate a relationship between your variables, by using relevant physical laws, principles like the input-output principle, your own mental picture of the processes involved, and any similar models you have met before, remembering always to simplify the problem as much as you can.
 - State what simplifying assumptions you have made in formulating your model.
3. Formulate the mathematical problem
 - Re-interpret your initial problem in terms of your model; convert the modelling problem into a mathematical one ...
4. Solve the mathematical problem
 - ... and solve it.
5. Interpret the solution
 - Decide what data you will need in order to specify the values of any parameters, and collect them.
 - Make any qualitative checks you can think of, to confirm that you are on the right track.
 - Interpret your mathematical solution so as to give a solution to the original problem in appropriate terms.
6. Compare with reality
 - Decide what data you will need to validate your model, and collect them.
 - Think of ways of improving your model, by relaxing one or more of your simplifying assumptions.

5 A modelling exercise for you to tackle

What shape should a tin can be? 'Cylindrical, of course,' I expect you say (but is it really 'of course'?). But given that tin cans are cylindrical, what shape of cylinder: short and fat, long and thin, or somewhere in between? I'm thinking of the standard kind of can in which soup and beans and other foodstuffs are preserved. Most of the ones in my kitchen cupboard contain around 400 grams of food and are rather taller than they are wide. There's a degree of uniformity in the general shape of cans which I find quite surprising, when there are so many different brands; even imported cans seem to be the same shape as home-manufactured ones. Is this tradition, or is it because they all conform to some ideal shape; and if the shape is not ideal, could there be advantages in changing it?

In this section you are asked to investigate the best shape for a tin can. This is a modelling problem for you to do, but you are not left entirely on your own to tackle it: the problem has been broken down into steps.

Exercise 1

Consider the best shape for a tin can to contain a given amount of baked beans.

- (i) It all depends what you mean by 'best' What do you think the word 'best' should mean, in the phrase 'the best shape for a tin can'?
- (ii) Try to formulate a clear statement of a modelling problem out of this. (A clear statement doesn't contain imprecise words like 'amount'!)
- (iii) Specify your variables.
- (iv) How does the volume constraint affect things?

If you don't worry too much about seams and joins, you should find it fairly straightforward to write down a formula which will help you to solve the problem.

- (v) Do so, expressing the area A in terms of radius r and volume V .
- (vi) List the assumptions that you made in deriving this formula.

It is worth pausing and checking that this formula makes sense.

- (vii) What happens to A as $r \rightarrow 0$, and as $r \rightarrow \infty$? Interpret your answers in terms of the shape of the can.
- (viii) Find the shape of a can of minimum area. Summarize your answer in words. Are tin cans really this shape?
- (ix) What have we left out?
- (x) Modify the formula for the area of tin-plate used to make a can so as to take into account your answer to part (ix). Solve the modified problem. Comment on your answer.
- (xi) Would square section cans use less tin-plate than is used in the optimal solution found in the previous part?

[Solution on page 51]

Appendix: Solutions to the exercises

Solutions to the exercises in Section 2

1. (i) The problem is to find how frequently it is necessary to give successive doses of a drug used in the treatment of a disease. The problem arises because the concentration of the drug in the patient's bloodstream decreases over time as the drug is excreted via the kidneys; the drug ceases to be effective if its concentration falls below a lower limit, but there is an upper limit to the size of a dose because the drug is toxic if its concentration is too high. The problem is important because the administration of the drug involves injection, which is painful, and so it is desirable to find how to administer it so as to ensure safety and effectiveness while avoiding unnecessarily frequent injections. To have a rule for this, in terms of relevant characteristics of the patient, would be better than simply waiting until the patient became ill again before repeating the injection.

It seems to me that, in this answer, I have been able to describe the essential features of the problem without naming the disease or the drug, and without having to discuss the details of the way it is administered. I think, therefore, that the specific problem considered in the report is just one of a general class concerned with the use of drugs in the treatment of a wide variety of diseases, and indeed with other situations such as the administration of painkillers. May there not even be situations entirely unconnected with medicine where similar processes are at work: what about the chlorination of water in a swimming pool?

(ii) The model describes how *the quantity of drug in the body* varies with *time*, for a given patient, as a consequence of excretion via the kidneys.

Or:

The model describes how *the concentration of drug in the bloodstream* varies with *time*, for a given patient, as a consequence of excretion via the kidneys.

The model does actually predict, in addition, how the quantity (or concentration) of drug at a fixed time depends on the patient's apparent volume of distribution, and on the size of the initial dose. However, we are primarily concerned with the time, since the question we have to answer is *how frequently* to repeat the dose; in the interests of simplicity we should concentrate on the time dependence first (and when we do so, the other factors enter naturally anyway).

The quantity of drug (and its concentration) will clearly decrease with time; moreover, the less drug there is, the less concentrated it will be, the less effective will the kidneys be in removing it, and so the more slowly it will decrease.

(iii) It is assumed that the drug 'is distributed uniformly throughout the bloodstream ... at all ... times' (assumption (b)). This means that, at any given time, the mass of drug contained in any sample of blood is directly proportional to the volume of the sample; the concentration is the constant of proportionality. To put it another way, the concentration of the drug is the mass of drug in a sample of blood divided by the volume of the sample; the assumption of uniformity says that this will not depend on the size of the sample nor on whereabouts in the bloodstream it is taken, but only on the time at which it is taken. 'Sample'

here must be taken in a hypothetical sense: I am describing what would happen if one did take a sample, rather than supposing that an actual sample is taken.

Concentration is analogous to density.

In a more complicated (and more realistic) model one might relax this assumption, and consider what happens if the drug is distributed unevenly through the bloodstream. It would become necessary to define concentration in a more sophisticated way. Concentration will depend on position in the circulatory system. It is defined as the mass of drug in a sample of blood taken from the neighbourhood of the given position, in the limit as the volume of the sample tends to zero.

When the drug is first administered it is concentrated close to the point at which it enters the bloodstream; as the blood circulates it carries the drug along with it, mixing it with the blood until it becomes effectively uniformly distributed. In the model developed here, that process is ignored, or if you prefer, regarded as taking place instantaneously. Note how fair the report is in pointing this out as an assumption, rather than leaving the reader to work it out.

(iv) There are two possible answers to this question, depending on how you answered part (ii). If you chose the quantity of drug, then your answer to this question should be

- q* the quantity – better, mass – of drug in the body, in milligrams,
- t* the time since the drug was administered, in hours.

If you chose the concentration, on the other hand, then your answer should be

- c* the concentration of drug in the bloodstream, in milligrams per litre,
- t* the time since the drug was administered, in hours.

The relationship between *q* and *c* requires a little thought. The equation $c(t) = q(t)/V$ appears to be a restatement of the definition of concentration given above (under the assumption of uniform distribution), but is not quite, because *q(t)* is defined as the quantity of drug in *the body*, not the bloodstream, and as pointed out in assumption (c), the drug is distributed to parts of the body other than just the bloodstream. Both variables, *q* and *c*, are required: *q* because its initial value is the size of the dose, and we can't tell how this is divided between the bloodstream and the rest of the body; *c* because it is the concentration which determines the effectiveness (and the toxicity) of the drug.

Note also that the specification of the time origin as the moment of administration of the drug depends on the assumptions that the conversion of aminophylline to theophylline, and the distribution of theophylline uniformly through the bloodstream, take place instantaneously (assumptions (a) and (b)).

As long as we are discussing a single patient, the quantity *V* (the apparent volume of distribution) should be considered as a constant rather than a variable; and as long as we are discussing what happens to a single dose, the size of the initial dose *D* should also be considered as a constant. Quantities like *V* and *D*, which are not fundamental variables, but whose values determine the conditions in any single instance of the model, are often called *parameters*.

(v) For my money, the best answer to this question is the differential equation relating q and t :

$$\frac{dq}{dt} = -\lambda q.$$

This is, after all, a direct mathematical description of the process of excretion of the drug via the kidneys. I suppose a case could be made for describing the formula for the concentration,

$$c = \frac{D}{V} e^{-\lambda t},$$

as a model, but the term is usually reserved for the most basic relationship between the variables arising directly from the assumptions, rather than those which may subsequently be derived by applying mathematical techniques, such as solving a differential equation.

The differential equation could have been reformulated in terms of c before any attempt was made to solve it, instead of solving it first and then substituting for q in terms of c . This would have given a model in terms of c and t :

$$\frac{dc}{dt} = -\lambda c.$$

(vi) The concentration is given by the formula

$$c = \frac{D}{V} e^{-\lambda t},$$

where D , V and λ are positive constants. The graph of c against t is thus a negative exponential, and looks like this.

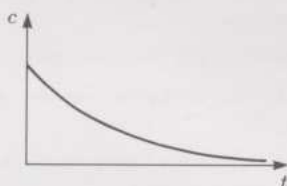


Figure 1

The evidence in the report that this is the correct formula for the concentration is based on the results of a (somewhat idealized sounding) experiment to measure the concentration of drug in blood samples taken at regular intervals. (Note that the description of the experiment states that time is allowed for the drug to diffuse throughout the bloodstream: we should not expect the experiment and the predictions of the model to match initially, while the real drug is not uniformly distributed, because this is not accounted for in the model.) If c depends exponentially on t , then the plot of $\log_e c$ against t should be a straight line, and the experimental results confirm this.

(vii) The question reduces to consideration of whether the equation

$$q(t + \delta t) - q(t) = -\lambda q(t) \delta t,$$

derived from the assumption by use of the input-output principle, holds exactly when $q(t) = D e^{-\lambda t}$, when δt is not necessarily small. Well, in this case

$$q(t + \delta t) - q(t) = D e^{-\lambda(t+\delta t)} - D e^{-\lambda t} = D e^{-\lambda t} (e^{-\lambda \delta t} - 1),$$

whereas

$$-\lambda q(t) \delta t = -\lambda D e^{-\lambda t} \delta t.$$

Thus

$$q(t + \delta t) - q(t) \neq -\lambda q(t) \delta t.$$

The assumption is modified in the course of finding the output, where it is pointed out that the proportionality of the output to the duration of the interval is only

approximately correct, though the approximation improves as δt decreases. (In the final report of this model, to be found in the *Guide to Mathematical Modelling*, the assumption will be corrected to reflect this modification.)

(viii) Yes, it does. Let the rate of flow of blood through the kidneys (which I assume to be constant) be ρ litres per hour. Then in a time interval of length δt hours, the volume of blood flowing through the kidneys is $\rho \delta t$ litres, and the mass of drug passing into the kidneys is thus approximately $c(t) \rho \delta t$ mg. (Note that this deduction holds only for short periods of time, so that the variation of $c(t)$ with time t can be ignored: hence the use of the word 'approximately'.)

The number of drug molecules in this quantity of drug is proportional to the mass of drug, and since each molecule has the same chance of being filtered out, a fixed proportion of this mass of drug is excreted. Thus

$$q(t + \delta t) - q(t) \simeq -k c(t) \rho \delta t,$$

where k is the constant of proportionality. But $c(t) = q(t)/V$, from which it follows that

$$q(t + \delta t) - q(t) \simeq -\lambda q(t) \delta t,$$

with $\lambda = -k\rho/V$.

(ix) In order to apply the formula for the concentration to any particular patient, one needs to know the values of D , V and λ . There is no difficulty about D , which is the dose: it is under the control of the doctor. But V and λ cannot be measured directly: to find V directly one would first have to drain off the patient's blood and measure its volume; and even that process, impractical as it is, would not provide the answer since V is the *apparent* volume of distribution. Hence the importance of the indirect method, in which the values of V and λ are deduced from the plot of $\log_e c(t)$ against t . This is a straight line, with slope $-\lambda$ and intercept $\log_e(D/V)$ on the vertical axis, as follows from the equation

$$\log_e c = \log_e \left(\frac{D}{V} \right) - \lambda t.$$

Since the value of D is known, the value of V can be calculated.

Even so, this is not a really satisfactory solution so far as clinical use is concerned. A seriously ill patient can't be subjected to the kind of experiment described in the report before the drug which will ameliorate his or her condition can be administered. There are two facts, additional to the model as discussed so far, which make it useful in drug therapy; both are mentioned explicitly in the report. First, there happens to be a simple relationship between apparent volume of distribution and body weight, so that the value of V can be deduced in a straightforward way from the result of weighing the patient. Secondly, the value of λ does not vary very much from patient to patient. Thus if we write

$$c(t) = \frac{D}{V} e^{-0.17t}, \quad V = \frac{1}{2}W,$$

we have a pair of formulae which can be used to find when the concentration drops below the therapeutic level and which requires only a knowledge of the size of dose and the patient's weight for its application.

(x) The coefficient 12 is simply the initial concentration of the drug (since it is $c(0)$, the value of c when $t = 0$). It is tempting, but unnecessary, to derive this by calculating D and V . In fact, the relationship between $c(0)$, D , V and W is much more useful as a method for calculating what dose is required to give a desired initial concentration (which must be kept below the toxicity level), rather than the other way round.

As t increases from 0, $c(t)$ decreases from 12. To find when the value of c drops to 5 we must solve the equation

$$5 = 12e^{-0.17t}$$

for t . The solution is

$$t = \frac{\log_e(12/5)}{0.17} \simeq \frac{0.88}{0.17} \simeq 5,$$

so the time for the concentration to fall to the lower limit of the therapeutic range is indeed about 5 hours.

(xi) There are three steps.

1. Find what dose is required to produce an initial concentration of the drug of 12 mg/l.
2. Find how long it will be before the concentration falls to 6 mg/l, when a further injection will be necessary.
3. Find what dose is required to bring the concentration up to 12 mg/l again. (A smaller dose will be required at this stage because even though the concentration has fallen to the bottom of the therapeutic range, there will still be a considerable amount of drug in the body.)

Further injections of the dose found in step 3 may then be repeated at the intervals found in step 2 for as long as the treatment is required.

The initial dose D mg is given by

$$D/V = 12, \quad \text{where } V = \frac{1}{2}W,$$

W being the patient's weight (in kg), so $D = 6W$.

To find the interval between further injections, we proceed in the same way as in the previous part, except that now we have to solve the equation

$$6 = 12e^{-0.17t}.$$

The solution is

$$t = \frac{\log_e 2}{0.17} \simeq \frac{0.69}{0.17} \simeq 4.$$

Thus after about 4 hours the concentration will have dropped to about 6 mg/l. To boost it up to 12 mg/l again requires just half of the initial dose, that is, $3W$ mg.

The instructions to the nursing staff might look as follows.

1. Weigh the patient on admission.
2. Give an initial dose of 60 mg for every 10 kg of the patient's weight.
3. Give a booster dose of half the initial dose every four hours.

This assumes that there is a sufficient safety margin in working well within the therapeutic range, as we have done, that we need consider the patient's weight only to the nearest 10 kilograms. It is certainly a good idea not to complicate nursing procedures unduly, by demanding that the dose be calculated correct to ten places of decimals, or that the period between injections be timed to the nearest second. Nursing procedures must be pretty well-conditioned to small changes in the data (to adapt the terminology of numerical analysis) to be practical, as anyone who has been in hospital will know.

(xii) The answers to previous parts of the question point out some of them. There is bound to be a fair degree of uncertainty in any model of a bodily function, and in particular the value of the constant λ will vary somewhat from patient to patient: someone who eliminates the drug

relatively slowly is in danger of overdosing if the upper limit is set too high. Moreover, the relation between apparent volume of distribution and body weight is not a precise law. It is unrealistic to expect that the doses will be repeated at exactly four-hourly intervals, however well run the hospital is. Indeed, the therapeutic limits themselves may differ between individuals.

However, there is one consideration which probably outweighs all these, and it relates to the assumption that the drug is instantaneously mixed uniformly through the bloodstream. All the concentrations calculated in the model depend on this assumption. Now there is clearly a period of time, immediately after injection, when the drug is not uniformly distributed but is concentrated in a small region of the bloodstream; and so initially the true concentration, in this region, is much higher than the calculated one. It follows that quite a considerable degree of erring on the side of caution is necessary if toxicity is to be avoided.

(xiii) The two which are suggested by the answers to the previous two parts of this question are as follows.

1. Carry out an 'error analysis', to find out how sensitive the conclusions drawn from the model are to variations in the value of λ , the relation between V and W , the measurement of W , and the time between injections.
2. Model the mixing process which takes place after the injection while the drug is being distributed through the bloodstream.

You may have thought of others.

I do not pretend that either of my suggestions is easy to carry out!

(xiv) The model will consist of a differential equation for the quantity of drug in the body as a function of time, as for the intermittent case. A new symbol must be introduced, to stand for the rate of infusion: suppose it to be I mg/hour.

The input-output principle may be applied; the version given in the report still applies to the excretion of the drug by the kidneys, but now the drug is being continuously replenished. In a time interval of length δt hours, an amount of drug $I\delta t$ mg is administered, i.e.

$$\text{input} = I\delta t.$$

Thus the input-output principle gives

$$q(t + \delta t) - q(t) \simeq I\delta t - \lambda q(t)\delta t,$$

whence

$$\frac{dq}{dt} = I - \lambda q.$$

Note that this differential equation has a constant particular solution $q = I/\lambda$; ideally, one would like to choose I so that the value of I/λ corresponds to a concentration which lies in the middle of the therapeutic range, since a suitable initial dose will then ensure that the concentration remains always at this level. In fact, the general solution of the differential equation is

$$q(t) = I/\lambda + ke^{-\lambda t}$$

for some constant k , from which it follows that $q(t)$ tends to I/λ in the long run; so this is a very good regime because it tends to settle down to the desired behaviour of its own accord, and the initial dose doesn't have to be very accurate, as long as it is below the toxic level!

2. (i) When a car skids, its tyres may leave marks on the road. The purpose of the model is to find a method of using the lengths of such skid marks to work out how fast the car was going when it began to skid. This information may be useful to the police after a road accident, when they are trying to find out what happened. (It would be interesting to know whether an estimate of the car's speed obtained like this would be acceptable in a court of law as reliable evidence in a prosecution for dangerous driving or a similar offence. To answer this question, both legal and mathematical knowledge would be required. A mathematical modeller would not be expected to have the necessary legal expertise; but he or she should be prepared to tackle the mathematical issues, of which the most important are the reliability and accuracy of the results.)

(ii) The test car provides data which are needed for the calculation, in addition to the length of the skid marks in the accident. These data allow a direct comparison to be made between the length of the skid marks in the accident, for which the speed of the car is unknown, and the length of skid marks made by a car whose speed at the onset of the skid is known. In effect, the data for the test car are used to work out the deceleration of the car involved in the accident while it is skidding. This is possible because the conditions under which the test takes place are similar to those for the accident, apart from the speed of the car. The state of the road surface and the car's tyres, and other circumstances of the accident, will have a significant effect on the skid. They will vary a lot from accident to accident. It is easier to reproduce the conditions in a test than it is to make allowance for them in a table.

(iii) The relationship required is one which gives the speed of the vehicle at the onset of the skid in terms of the length of the skid marks. The longer the skid marks, the faster the car must have been going when it began to skid (all other things being equal), so the speed will be an increasing function of the length of the skid marks.

(iv) It sounds as though 'retardation' should mean the same as 'deceleration'. However, in the middle of the passage the phrase 'the above retardation' is used to refer to the formula $a = -u_t^2/(2s_t)$, so that retardation is equated with acceleration. It is therefore negative in this case. (Much confusion can be avoided by sticking to the term 'acceleration', and remembering that for a point moving in the positive direction its acceleration is positive when its speed is increasing, negative when its speed is decreasing.)

(v) The fundamental variables are the original car's initial speed, u , and the length of the skid, s . The data, which are the corresponding quantities for the test car, are represented by the symbols u_t (initial speed) and s_t (length of skid). All the symbols appearing in the final formula have now been identified, but there are still others, which occur only in the course of the calculations. These are a , the acceleration, and v the final speed, of either car. The final speed is zero in each case (so the symbol v is not really necessary). The reasons why the two cars may be assumed to have the same acceleration will be considered in parts (viii) and (ix). The question of units (or rather, the lack of them) is dealt with in the next part of the exercise.

(vi) No, it doesn't matter in this instance. One way of seeing why involves rewriting the final formula as

$$\frac{u}{u_t} = \sqrt{\frac{s}{s_t}}.$$

The left-hand side of this equation is the *ratio* of two speeds, while the right-hand side is the square root of the *ratio* of two distances. Now the ratio of two speeds is simply a number, and has the same value regardless of whether the speeds are measured in ms^{-1} , mph or leagues per century (provided that both speeds are measured in the same units). Likewise, the ratio of two distances takes the same value whatever unit of measurement is used. The omission of units in the description of the variables is therefore not important in this case. Some units must be used when one is making a measurement, of course; but provided the same units are used for the speeds, and the same units for the distances, it does not matter which. One could even use different systems of units for the two kinds of measurements: the formula will give the speed of the original car in mph if the speed of the test car is expressed in mph, even if the distances are measured in metres. However, mixing units is a practice which should be avoided unless you are really sure you know what you are doing.

(As you will know if you have taken *TM282*, the omission of units is not an oversight. In that course the approach to units of measurement is different from that taken in *MST204*, and in some ways is more sophisticated. It is based on the fact that any useful and significant formula takes the same form in any system of units, provided that consistent units are used throughout. Thus the formula relating initial and final speed, acceleration and distance for a point moving with constant acceleration is the same whether SI or imperial units are used. We therefore do not need to specify the units for the variables v , u , a and s in the formula $v^2 = u^2 + 2as$, provided we use consistent units for distance, speed and acceleration.)

(vii) The model which underlies the whole discussion is the standard model of the motion of a point moving in a straight line with constant acceleration, which you met in *Unit 4*, Section 1. The formula required here is the one relating the initial and final speed, acceleration and distance. In the box on page 15 of *Unit 4* this is given as

$$v^2 = v_0^2 + 2a_0(x - x_0),$$

which becomes the same as the equation used here when u is substituted for v_0 , a for a_0 and s for $x - x_0$.

(viii) The basic assumptions are as follows.

1. A skidding car has constant acceleration.
2. The test car has the same acceleration as the original car.

(ix) The support I would give for the two assumptions is as follows.

1. The text says that '... the frictional forces between surfaces are not dependent upon the speed of the car, only upon the mass of the car, the condition of the road and the type of surface'. This implies that the frictional force is constant. If the road is horizontal, it follows immediately that the acceleration will also be constant (assuming that friction is the only force acting). If the road slopes at a constant angle, the same will be true, since the additional component of force due to gravity will also be constant.
2. The next sentence says that the 'frictional force is assumed to be proportional to the mass of the vehicle, as are any ... forces due to gravity if the car is going up or down a hill'. The fact that the test car has similar tyres to the original car, and the conditions of the test

being similar to those of the crash, mean that the other factors affecting the frictional force are the same in the two instances. The total force component acting is therefore a constant times the mass of the car, where the constant is the same for the test car and the original car. Since acceleration is force divided by mass, this justifies the assumption that the two cars have the same acceleration.

(x) The rather wordy solution to the previous part of this exercise does suggest that a diagram will clarify what is going on. For an accident which takes place on a horizontal road, assuming that air resistance can be ignored, the force diagram looks like Figure 2.

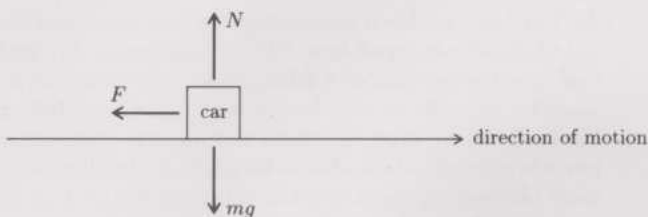


Figure 2

Here N denotes the magnitude of the normal reaction and F the magnitude of the frictional force. By Newton's second law,

$$ma = -F,$$

where m is the car's mass and a its acceleration in the direction of motion. Furthermore, since there is no vertical motion,

$$N = mg.$$

Since the car is skidding (with its wheels locked) the rule relating frictional and normal reactions, given in the box on page 30 of *Unit 15*, applies. This gives

$$F = \mu' N,$$

where μ' is the coefficient of kinetic friction. So

$$a = -\mu' g.$$

This says that the acceleration of a skidding car is constant, and that the magnitude of the acceleration is independent of the mass of the car and will be the same under all conditions leading to the same value of the coefficient of kinetic friction. The test is designed with the precise aim of producing conditions leading to the same value of the coefficient of kinetic friction.

(xi) The model will apply to a crash which takes place on a sloping road, provided the slope doesn't vary. Under these conditions a skidding car will have constant acceleration, and the magnitude of the acceleration will depend only on the coefficient of kinetic friction and the slope of the road.

The force diagram for a downhill skid is shown in Figure 3.

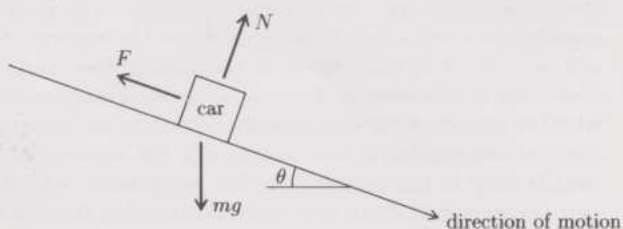


Figure 3

The equations of motion are

$$ma = mg \sin \theta - F,$$

$$N = mg \cos \theta,$$

$$F = \mu' N,$$

where θ is the angle of the slope. These equations lead to

$$a = (\sin \theta - \mu' \cos \theta)g.$$

As before, the acceleration is constant and the same for the test car as for the crashed car, provided that μ' is the same for each and that the test is carried out on road with the same slope as the road where the crash took place. (We must suppose that $\tan \theta < \mu'$, otherwise the car will never stop!)

The model will apply to a Rolls Royce provided that the test car has similar tyres. The size and weight of the crashed car (let alone the price) are not relevant, so long as the coefficient of kinetic friction can be duplicated in the test.

A crash into a headwind will not be covered by the model, because the headwind will introduce an additional force, of air resistance, which will depend on the speed of the car relative to the wind; so the assumption of constant acceleration will no longer apply.

If the road is wet at the time of the crash, but dries before the test, the model will not apply because the assumption that the acceleration is the same in the test and the crash will not be valid: the coefficient of kinetic friction for a wet road will be different from that for a dry road.

(xii) The formula $u = \sqrt{-2as}$ does predict that u is an increasing function of s , as expected. However, by itself it is not enough to solve the problem, because we have no direct way of finding the value of a . The test gets round this difficulty, and provides a value of a , so long as the assumptions that have just been discussed are satisfied.

(xiii) The symbol s is being used with two different meanings: as the distance travelled in general by a point moving with constant acceleration, and as the length of the skid marks of the crashed car. The method used in the passage to analyse the skids is to take the general formula for motion with constant acceleration in the form $v^2 = u^2 + 2as$ and substitute the particular values of v , u , a and s appropriate to the particular circumstances. In the case of the crashed car we must substitute the length of the skid marks of the crash for s ; since the same symbol has been used for this as for the general distance, we are led to the odd looking equation $s = s$. The moral of this story is that it is better to observe the rule that each symbol should have only one meaning.

(xiv) If the skid marks of the crashed car are four times as long as those of the test car, then $s/s_t = 4$, so $u = 2u_t$, which means that the crashed car was travelling at 80 mph when it began to skid, and was exceeding the 70 mph speed limit.

(xv) The instructions to a policeman going to the scene of an accident for the first time might take the following form.

1. If there has been a skid which has left a mark on the road, measure the length of this mark.
2. Check that the tyres on the crashed car are in a comparable state of wear to those on the patrol car, and that the circumstances have not changed since the accident in a way which will affect the slipperiness of the road (for example, that the weather conditions have not changed).

3. With a clear road, driving the patrol car, approach (at a safe speed) the spot where the accident took place, and induce a skid. Note the initial speed of the patrol car, and measure the length of its skid.
4. To calculate the speed of the crashed car when it began to skid, divide the length of the crash skid by the length of the test skid, take the square root of the result, and multiply it by the speed of the patrol car (in mph). The answer will be an estimate of the speed, in mph, of the other car when it started skidding.

(xvi) The model has been based on the assumption that $v = 0$. If this is not the case, then, provided that the test still gives the correct value for the acceleration, we have

$$u = \sqrt{v^2 - 2as} > \sqrt{-2as}.$$

So the actual speed of the crashed car would have been greater than that estimated using the model.

Solutions to the exercises in Section 3

1. The actual problem described in the summary of the model of double glazing is to find how the rate of transfer of heat through a double-glazed window depends on the difference between the air temperatures on either side of it. For many applications it would also be of interest to find how the rate of transfer of heat depends on the design of the window – for example, on the width of the space between the two panes, or the thickness of glass used. This information is contained in the model, but it is necessary to know how the U value depends on such features to obtain it explicitly.

The most likely use of a model of heat transfer through double glazing is in calculations concerned with saving energy. For example, to calculate the savings resulting from installing double glazing in one's home, one would have to estimate the annual heat loss through the windows with and without double glazing; the difference would be the amount of energy saved per year, and the cost of providing this energy would be the financial saving. This would have to be compared with the cost of installing the double glazing. Again, any attempt to estimate the reduction in national demand for energy resulting from a policy of improving the insulation of buildings would have to take into account the effects of double glazing, among the various forms of insulation.

2. The shot-putter's objective is to put the shot as far as possible. How far he or she can put it depends in part on physique and strength, in part on exactly how the shot is thrown. A mathematical model of the motion of the shot cannot help much with the former, but it can help with the latter. One important factor affecting the range of the shot is the angle at which it is launched. The formula relating the range to the angle of launch can be used to find which angle gives the maximum range. The shot-putter can train to use this launch angle, and so try to ensure that he or she achieves the maximum range.

3. The sandwich shop manager will want to order just sufficiently many sandwiches to satisfy one day's demand; ordering too many sandwiches will lead to waste, while ordering too few will result in disappointed customers. The problem therefore is to predict demand. We must next ask what factors will affect the demand for sandwiches on any

particular day. There are several factors which will affect the overall level of demand, such as the location of the shop; however, the most important question, so far as the manager is concerned, is how the demand varies from day to day.

There are three obvious factors affecting the daily variation in demand: what day of the week it is, what time of the year it is, and what the weather is like. So the basic problem is to find out how the daily demand for sandwiches will depend on the day of the week, the time of year and the weather.

The distinctive feature of sandwiches is their short shelf-life. Another commodity with the same feature is newspapers. The demand for newspapers will depend on different factors; one major factor will be what newsworthy events are taking place, if any.

4. The basic problem underlying the analysis of the time of fall of objects dropped from Clifton Suspension Bridge is to find how the *position* of a falling object (modelled as a particle) depends on *how long a time* it has been falling. We begin by selecting an origin and an axis, and we use the particle's coordinate to define its position. In this case the most obvious choice of origin is the point from which the particle is dropped, and the most obvious axis points vertically downwards. It is also sensible to measure time from the moment the particle is released. The main variables are therefore

- x , the particle's coordinate with respect to the chosen axis (x is the distance the particle has fallen, in metres),
- t , the time in seconds since the particle was released.

It is tempting to consider the height of the bridge as a 'variable'. We have used quotation marks here to point out the oddity of this suggestion: the effects of any significant variation in the height of a bridge do not bear thinking about. By widening the problem to cover the whole period of the motion, and by using the particle's coordinate as a variable, we avoid this absurdity.

Position and time are very often the main variables in mechanics problems.

5. In modelling the repayment of a mortgage, the basic problem is to find how the *amount of money still owed* depends on *how long* one has been repaying the mortgage. The main variables are therefore

- u_r , the amount of money owed after r years, in £,
- r , the number of years since the mortgage was taken out.

6. This problem is concerned with deciding the *position* of a supermarket: so one (or more) of the variables must specify the position. One way of doing this which would be appropriate to this kind of problem would be to use a map reference; in effect, a map reference is a pair of coordinates, so we could say that the coordinates of the possible sites of the supermarket are a pair of main variables. The next question is to ask what determines where the supermarket will be sited. The basic factor is presumably the number of customers it will attract. Leaving aside such matters as whether people prefer one supermarket chain to another, the simplest assumption is that people who use supermarkets usually shop in the nearest one. For any possible site of the new supermarket, there is an area surrounding the site with the property that, if you live in this area, and if the new supermarket were built at that site, the supermarket closest

to where you live would be the new one. This area we call the catchment area of the site. The crucial question is *how many people live in the catchment area*. The management of the supermarket chain will probably consider only sites for new supermarkets for which the catchment area contains a certain minimum number of people; and it will seek to choose the site to maximize the number of people living in the catchment area. The main variables are therefore

- the number of people living in the catchment area of any possible site of the new supermarket,
 - the coordinates of each possible site.
7. The main simplifying assumptions are
- the object may be treated as a particle (as in the solution to Exercise 4 of this section),
 - the particle is dropped vertically and moves in a vertical straight line,
 - (for a first model, at least) air resistance can be ignored, so that the only force acting on the particle is its weight.

8. The fundamental assumption is that the system (the double-glazed window and its surroundings) is in a steady state: that is, neither the internal or external temperature, nor the rate of heat transfer, is changing in time. In addition we are assuming that the heat is transferred in one direction only (perpendicularly to the plane of the window), so that the effects of the edge of the window are ignored; and we are ignoring the effects of frames and glazing bars. The window is assumed to be closed and draught free. We are also assuming standard models of conduction and convection.

9. For a first model it would be wise to make as many simplifications as possible, consistent with retaining some degree of realism. Take the windscreen to be flat and rectangular. Deal with a single wiper only to begin with (that is, solve the problem for the rear window first, or assume the car is a Citroën). Assume that the area swept out by the wiper blade has the shape shown in Figure 1.

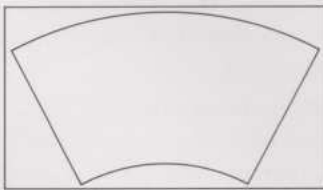


Figure 1

The curves are sectors of circles with the same centre, and the straight sides are segments of radii of the larger circle. Such a shape should probably be called a sector of an annulus. Finally, assume that the wiper blade never lifts off the glass, so that the region swept out lies entirely within the rectangular area of the windscreen.

(The problem we are left with after these simplifications is still quite hard enough: it is to find the largest sector of an annulus which can be fitted inside a given rectangle. A full solution would require, in effect, the specification of the dimensions of the wiper: length of arm, length of blade and angle of sweep; and the location of the point about which it rotates.)

10. In each year the building society (or whoever is the lender) will charge interest on the amount still owing at the beginning of the year; this will be added to the outstanding amount. During the year, twelve equal monthly repayments are made. When these are subtracted in total from the amount outstanding plus interest, what remains is the amount owed at the beginning of the next year.

If the amount owed at the beginning of year r is u_r and the interest rate is $I\%$ per annum, then the interest to be added is $Iu_r/100$, while the total repayment in the year is $12M$ if the monthly repayment is M . Thus

$$u_{r+1} = \left(1 + \frac{I}{100}\right) u_r - 12M.$$

In the example in *Unit 1*, the interest rate is 15%, so the formula becomes

$$u_{r+1} = 1.15u_r - 12M.$$

(All variables representing sums of money are in units of £, whereas I is an annual percentage.)

11. The force diagram is shown below.

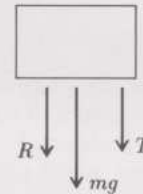


Figure 2

The forces acting on the car body are the tension in the spring, T , and the resistance due to the damper, R .

In order to derive the equation of motion, we must first define the coordinate. We shall take the x -axis to point upwards, with the origin at the equilibrium position of the car body (there are other possibilities, of course). The equation of motion is obtained by using Newton's second law; in terms of the symbols for the forces shown in the figure we have

$$m\ddot{x} = -R - mg - T.$$

We must next express R and T in terms of x and $\dot{x}$. The resistance force, assumed to be due to linear damping, is given by $R = r\dot{x}$, where r is a constant. The expression for the tension may be derived quickly as follows. For a perfect spring the tension will be a linear function of x . In other words, we must have $T = kx + l$, where k is the stiffness of the spring and l is an as yet unknown constant. We may find l by noting that $T + mg = 0$ when $x = 0$, which corresponds to equilibrium. Thus $l = -mg$, and $T = kx - mg$. The equation of motion is therefore

$$m\ddot{x} = -r\dot{x} - mg - (kx - mg),$$

which may be rewritten

$$m\ddot{x} + r\dot{x} + kx = 0.$$

12. To express t in terms of x we must work out the time taken to find a parking space, and the time taken to walk from there to the shops, and add the two together. The time taken to find a parking space is assumed to be inversely proportional to x , so it can be written k/x where k is a constant. I then have to walk a distance x at a speed of (say) v , which takes a time x/v . The total time t is therefore

given by

$$t = \frac{k}{x} + \frac{x}{v}.$$

All quantities here are measured in appropriate SI units – or, indeed, in any consistent system of units (yards and minutes might be considered more suitable than metres and seconds). The constant of proportionality k must have units of m s (in the SI system).

The question asks only for the formulation of the problem, but it is easy to find the value of x which minimizes t , and it turns out to be $\sqrt{kv}$. (Note that the units make sense: since k is in m s and v in m s^{-1} , the result is in metres as required.) I know my own walking speed, or can easily find it; the difficulty with this model will be to estimate k . One reason for this is that although it is stated that k is a constant, it will actually vary from day to day, and even during the day. In fact, the value of k will depend on how much competition there is for parking places, that is, on how many other people want to go shopping at the same time as I do.

13. The solution to the logistic equation,

$$P = \frac{M}{1 + (M/P_0 - 1)e^{-at}},$$

has the following features that suggest that it would be a satisfactory model of the growth of a population. In the first place, P is an increasing function of t (provided that $P_0 < M$), so it does indeed model population growth. (In fact, the same general formula would model a growing, stable or declining population, according as $P_0 < M$, $P_0 = M$ or $P_0 > M$.) Secondly, there is an upper limit to P , namely M , and P approaches M more and more closely as $t \rightarrow \infty$. This accords with the expectation that a population cannot increase indefinitely, but that there is a natural upper limit to its growth.

Thirdly, if P_0 (the initial size of the population) is small compared with M , then for small values of t , $P \simeq P_0 e^{at}$. (This result is shown in detail on page 21 of *Unit 3*.) This means that a population which is small initially will grow exponentially to begin with. This is the expected growth pattern of a population while there is no competition for food and space. These features of the solution of the logistic equation are immediately apparent from its graph (shown below); indeed, sketching the graph would in many ways have been the best way of answering the question.

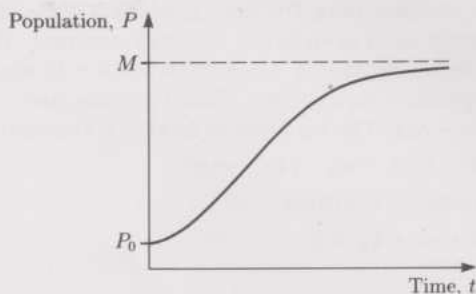


Figure 3

There are two ways of rearranging the formula to provide a way of checking the model against data. The first is to rewrite it as

$$\left(\frac{M}{P_0} - 1\right) e^{-at} = \left(\frac{M}{P} - 1\right),$$

and then take logs to obtain

$$\log_e \left(\frac{M}{P_0} - 1\right) - at = \log_e \left(\frac{M}{P} - 1\right).$$

This says that a plot of $\log_e(M/P - 1)$ against t should give a straight line. In order to use this method, however, one must first estimate M , the size of the equilibrium population; this will not be easy if the population is still growing. There is an alternative method, which does not have this drawback, which can be used if the data are for the population sizes at equal time intervals. Let P_n be the population when $t = nT$, where T is some fixed length of time (one year, for example). Then

$$\begin{aligned} \left(\frac{M}{P_{n+1}} - 1\right) &= \left(\frac{M}{P_0} - 1\right) e^{-a(n+1)T} \\ &= \left(\frac{M}{P_0} - 1\right) e^{-anT} e^{-aT} = \left(\frac{M}{P_n} - 1\right) e^{-aT}, \end{aligned}$$

from which we obtain

$$\frac{P_{n+1} - P_n}{P_n} = \left(\frac{1 - e^{aT}}{M}\right) P_{n+1} - (e^{aT} - 1).$$

It follows that a plot of $(P_{n+1} - P_n)/P_n$ against P_{n+1} should produce a straight line. (See Subsection 2.4 of *Unit 3* for more details of these methods.)

(The logistic equation is not a satisfactory model for the growth of every population, of course. It takes no account of population movements, of the interaction of the population with its environment (overgrazing, the effects of predators), or of random influences like the weather, to mention just three of many factors which could affect population growth. Populations of species for whose development these are important factors will not follow the logistic pattern of growth. Nevertheless, the logistic equation is a very useful basic model in population dynamics, much as motion in one dimension under gravity is useful in particle dynamics despite the factors – like air resistance – it leaves out of account.)

14. The general solution of the recurrence relation for the mortgage problem (with interest at 15%) is

$$u_n = B(1.15)^n + 80M.$$

The initial condition, that $u_0 = 20\,000$, can be used to determine the constant B in terms of M , giving

$$B = 20\,000 - 80M.$$

We wish to determine the monthly repayment required to pay off the loan after 25 years: this leads to the equation

$$u_{25} = 0 = (20\,000 - 80M)(1.15)^{25} + 80M,$$

or

$$M = \frac{20\,000(1.15)^{25}}{80((1.15)^{25} - 1)} = 257.8324\dots$$

(These results are derived in detail on page 36 of *Unit 1*.)

The practical problems of implementing the solution arise because the calculation of u_{25} is ill-conditioned with respect to small changes in the value of M ; the monthly repayment is a sum of money, and so the exact value of M cannot be used, but must be rounded. As noted in the description of the model, rounding M up to 258, and working throughout to the nearest whole number of pounds, results in an overpayment of £427. If, on the other hand, one works throughout to the nearest penny, taking the monthly repayment to be £257.83, then after 25 years the sum of £5.99 will still be owing. This is despite the fact that the 'error' in each repayment is 0.24p, which over 25 years accumulates to only 72p.

In fact, mortgage calculations are very sensitive to changes in almost any of the relevant data. One good illustration of this is the fact that when the interest rate is 15% and the loan is £20 000, a monthly repayment of £250 is required simply to pay the interest. In other words, if the monthly repayment is £250 then the amount owed stays constant at £20 000. But if the monthly repayment is £258 the loan is more than paid off after 25 years; while if the monthly repayment is £249 then the amount owed will actually increase.

The moral of this exercise is that it is often not enough just to calculate a numerical answer to a modelling problem. One should also consider the sensitivity of the answer to the accuracy of the data, and the effects of any approximations used.

15. First, one can check that the formula makes sense. Note that t is an increasing function of m , and that $t = 0$ when $m = 0$. The graph of $km^{2/3}$ is shown in Figure 4.

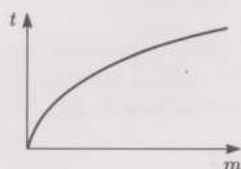


Figure 4

The formula therefore predicts that as the mass of the piece of food to be cooked (a potato, for example) increases, the cooking time also increases; but that twice the mass of food doesn't need twice the cooking time, but only about 60% longer ($2^{2/3} = \sqrt[3]{4} \approx 1.59$). These predictions seem qualitatively correct.

There are, in principle, two ways of checking the formula in more detail. One is to carry out some experiments. The other is to find a theoretical derivation of the formula.

To check the formula by experiment, you would need to cook a number of samples of the same food but of different masses. In order to tell when the food was cooked you would need a cooking thermometer or temperature probe of some kind, which could be stuck into the sample being cooked and which would register the temperature at its centre. Each sample would have to be cooked by itself, the oven temperature being the same for each.

You would record how long it took for the temperature of each sample to reach some predetermined value. When the experiment was complete, you would plot the logarithm of the cooking time against the logarithm of the mass for all the samples, and expect to obtain a straight line of slope $\frac{2}{3}$.

The concepts of *Unit 12* could be used to derive the formula theoretically, though the problem is much more complicated than any considered in that unit. Even so, we can note that the main heat transfer process involved in the heating of the food is conduction, which suggests that the value of the constant of proportionality k will depend on the thermal conductivity of the food, and also the oven temperature. It is also worth remarking that the rate of transfer of heat into the food will be proportional to the area of its surface; the surface area of a sphere is proportional to the square of its radius, and therefore to the $2/3$ power of its volume, or

equally the $2/3$ power of its mass (supposing the density to be uniform). This is supportive evidence for the formula, though not of course a conclusive derivation. However, this approach does suggest that the shape of the object being cooked is not a significant factor, because the relation between surface area and volume just derived holds true, in a sense, for any shape. This remark is based on the observation that a large baking potato (say) is much the same shape as a small potato; the larger one is simply a uniformly scaled up version of the smaller one. But when you uniformly scale an object by a factor λ , you scale its surface area by a factor λ^2 , and its volume by a factor λ^3 . This means that, for any collection of similar objects (potatoes, for example), the surface area of any one of them is proportional to the $2/3$ power of its volume, or the $2/3$ power of its mass. If the hypothesis that cooking time is proportional to surface area is correct, then the cooking time for one of these objects is indeed proportional to $m^{2/3}$ where m is its mass.

To find how long it would take to bake a potato, one would need to know the value of the relevant constant k . This will take the same value for all potatoes cooked at a given oven temperature, but will depend on the oven temperature chosen. The choice of oven temperature has a lot to do with the eating quality of the food when it is cooked. You must select an oven temperature that produces baked potatoes as you like them; it will then be sufficient to carry out one test baking to determine the appropriate value of k . Armed with this information – which, presumably, you will have recorded somewhere for future use – you need only weigh your potato to work out the cooking time. In order to avoid disappointment, it would be wise when recording the value of k to make a note of the system of units you have adopted, so that you will know whether to weigh the potato in grams or ounces. This is a situation in which SI units are not the most useful, since cooking times are rarely measured in seconds.

(Cookery books – that is, books of recipes – tend to avoid such sophisticated matters as $2/3$ powers. I know of only one cookery book in which the $2/3$ power formula is discussed, and that is an old one: *The Grammar of Cookery* by Philip Harben, Penguin 1965. Harben claims that the formula 'holds good for every time calculation in cookery, be it roasting, boiling eggs, baking cakes, deep-frying – every single thing'. But he doesn't expect the reader to use the formula to work out the cooking times: he gives a table of cooking times in each recipe.)

16. The exponential model is fairly accurate for many populations in a state of rapid increase, over relatively short periods of time. Its major drawback is that it predicts unrestricted and indefinite growth, which cannot continue in fact since resources are finite.

In the exponential model it is assumed that the rate of change in the population is a constant multiple of the population size; in other words, that the proportionate growth rate of the population is constant. (The proportionate growth rate of the population is the difference between the proportionate birth rate and the proportionate death rate. The assumption that each of these is separately constant will lead to the exponential model, but in fact it holds in more general circumstances.)

To model competition for scarce resources we assume that the larger the population is, the smaller its rate of reproduction will be, or the larger its mortality rate, or both. This amounts to saying that the proportionate growth rate of the population is not constant, but depends on the population size; and that it is a decreasing function of population size. In the absence of any detailed information about exactly how the proportionate growth rate $g(P)$ varies with the population size P , we make the simplest possible assumption, which is that it is a decreasing linear function of P : we may thus write

$$g(P) = a \left(1 - \frac{P}{M}\right),$$

where a and M are positive constants. This particular form is chosen because the constants a and M are relatively easy to interpret in terms of the model. It follows from the assumptions that the population size P must satisfy the differential equation

$$\frac{1}{P} \frac{dP}{dt} = a \left(1 - \frac{P}{M}\right),$$

which can be written

$$\frac{dP}{dt} = aP \left(1 - \frac{P}{M}\right).$$

This is the logistic equation. We see that M is the value of P for which $dP/dt = 0$, which means that a population of this size will remain at the same size indefinitely; so M is the equilibrium population size. When P is small the differential equation approximates to the exponential equation

$$\frac{dP}{dt} = aP;$$

this shows that the logistic model incorporates the exponential model as an approximation for small populations, and that a is the proportionate growth rate when the population is going through this approximately exponential phase of growth.

17. In the Clifton Suspension Bridge problem, the result of introducing the resistive force, assumed proportional to speed, is to change the equation of motion from

$$m \frac{dv}{dt} = mg$$

to

$$m \frac{dv}{dt} = mg - kv,$$

where k is a positive constant. The qualitative effect of this change on the predicted motion is quite significant. For example, the simple model predicts that the speed of the falling body will increase linearly with time, and that in principle the speed could become arbitrarily large. But when the resistive force is included, the speed can never exceed the terminal speed $v_T = mg/k$. On solving the equation of motion for v in this case, we find that

$$v = \frac{mg}{k} (1 - e^{-kt/m}) = v_T (1 - e^{-kt/m}).$$

(See page 42 of *Unit 4* for the full derivation of this result.) It follows that when t is large, so that $e^{-kt/m}$ is small, $v \simeq v_T$. When t is small, on the other hand, $e^{-kt/m} \simeq 1 - kt/m$ (using the first two terms in the Taylor series of the exponential, which is given in Subsection 6.4 of the *Handbook*); thus

$$v \simeq \frac{mg}{k} (1 - (1 - kt/m)) = gt,$$

which agrees with the result in the absence of the resistive force.

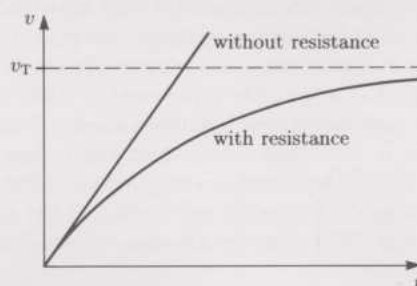


Figure 5

The predicted time of fall will be longer with the resistive force than without, since the speed is smaller. The relation between distance fallen and time is obtained by solving the differential equation

$$\frac{dx}{dt} = v_T (1 - e^{-kt/m});$$

the solution turns out to be

$$x = v_T t - \frac{mv_T}{k} (1 - e^{-kt/m}).$$

Again, it is interesting to consider what happens for large and for small t . For large t the term $e^{-kt/m}$ is small, and the motion is given approximately by

$$x \simeq v_T \left(t - \frac{m}{k}\right),$$

which is that of a body falling with constant speed v_T . For small t we have $e^{-kt/m} \simeq 1 - kt/m + \frac{1}{2}(kt/m)^2$, whence

$$\begin{aligned} x &\simeq v_T t - \frac{mv_T}{k} \left(1 - \left(1 - kt/m + \frac{1}{2}(kt/m)^2\right)\right) \\ &= \frac{1}{2}gt^2. \end{aligned}$$

Once again, we obtain the result for the motion under gravity alone as an approximation for small t (though we have to include an extra term in the Taylor expansion of the exponential).

To find how long it takes to fall a specific distance h , we would have to solve the equation

$$h = v_T t - \frac{mv_T}{k} (1 - e^{-kt/m})$$

for t . Unfortunately, it is not possible to solve this equation explicitly to give t in terms of h . However, a sketch of the graph of x against t shows that the equation has a solution lying between $\sqrt{2gh}$ and $h/v_T + m/k$.

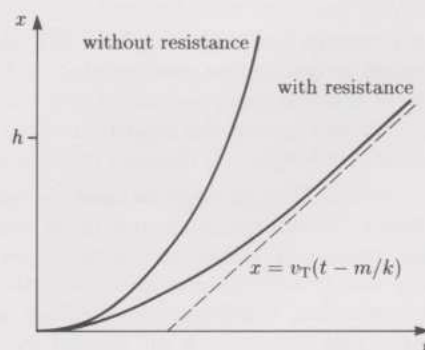


Figure 6

It is also worth noting that one consequence of introducing the resistive force is that the results depend on the particle's mass. For motion under gravity alone, this is not so: all bodies fall equally fast (as Galileo is supposed to have demonstrated in Pisa). But when air resistance takes effect,

two bodies of the same dimensions, one made of lead and one of gossamer, will have different times of fall.

18. The effect of taking the drug orally is that the drug in the bloodstream will be continually replenished as what remains in the stomach is absorbed. This is similar to intravenous infusion, but with one important difference: the stomach will not pass the drug to the bloodstream at a constant rate. It seems likely that the absorption of the drug in the stomach works something like the elimination process in the original model: the rate at which the drug is absorbed (and therefore transferred to the bloodstream) is proportional to the amount remaining in the stomach. This implies that the amount of drug in the stomach decreases exponentially, so we can write

$$Q(t) = De^{-\mu t},$$

where $Q(t)$ is the quantity of drug in the stomach, in mg, t seconds after the dose was taken. Here $D = Q(0)$ is the initial dose, again in mg, and μ is a positive constant which measures the rate of absorption. Now drug lost from the stomach is gained by the bloodstream, so that the rate of gain of drug by the bloodstream is $-dQ/dt = \mu De^{-\mu t}$. The equation for the quantity of drug in the bloodstream becomes

$$\frac{dq}{dt} = -\lambda q + \mu De^{-\mu t},$$

the first term on the right-hand side representing excretion via the kidneys (as before).

This equation can be solved readily by the integrating factor method (for example). The general solution, for $\lambda \neq \mu$, is

$$q = Ae^{-\lambda t} + \frac{\mu D}{\lambda - \mu} e^{-\mu t};$$

the solution which satisfies the initial condition $q(0) = 0$ is

$$q = \frac{\mu D}{\lambda - \mu} (e^{-\mu t} - e^{-\lambda t}).$$

The drug concentration is therefore

$$c = \frac{\mu D}{V(\lambda - \mu)} (e^{-\mu t} - e^{-\lambda t}),$$

where V is the apparent volume of distribution.

The graph of this function looks like Figure 7.

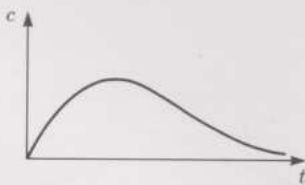


Figure 7

The concentration is initially zero, grows to begin with at the rate $\mu D/V$, reaches a maximum when

$$t = \frac{\log_e \lambda - \log_e \mu}{\lambda - \mu},$$

and then tails off towards zero again.

Before we could make any quantitative statements about the size and frequency of the dose in this case, we would have to know the value of the constant μ , which is a measure of the rate at which the drug is transferred from the stomach to the bloodstream. However, we can see how to proceed once this is known. The initial dose D must be chosen so that the maximum value of the drug concentration lies below the level at which toxicity occurs. It will take some time before

the concentration is sufficient for there to be a therapeutic effect; and second and subsequent doses must be administered before the concentration falls below the therapeutic range, in order to allow for this delay.

Solutions to the exercises in Section 4

1. The process to be modelled is the warming of something cold – a liquid contained in a bottle – when it is placed in surroundings which are at a higher temperature than it is. The problem, therefore, is to find how the temperature of the liquid varies with time. Let t be the time, in seconds, since the bottle was placed in the warmer surroundings; and let $\theta(t)$ be the temperature of the liquid, in degrees Centigrade, at time t . One would expect θ to be an increasing function of t , and to approach the temperature of the surroundings more and more closely as t increases but never to quite reach it.

The liquid's temperature increases because heat energy is continually being transferred to it from its surroundings; the most significant mechanism of heat transfer in this case is likely to be conduction through the wall of the bottle. To model this it will be necessary to make some assumption about the shape of the bottle: the simplest assumption which is at all realistic (in most cases) is that the bottle can be regarded as a cylinder. The heat conducted through the wall of the bottle will in fact warm both the liquid and the bottle; to simplify matters, the warming of the bottle itself will be ignored, so it will be assumed that all of the heat warms the liquid. It will also be assumed that the temperature of the surroundings remains constant as the liquid warms up. The effects of convection will be ignored, as will the opening of the top of the bottle.

In the steady state, the rate of transfer of heat by conduction through the side and ends of a cylinder of length l metres, radius r metres and thickness b metres is q watts, where

$$q = \kappa \left\{ \frac{2\pi l}{\log_e((r+b)/r)} + 2\frac{\pi r^2}{b} \right\} \Delta\theta,$$

where κ is the thermal conductivity, in $\text{W m}^{-1} \text{ } ^\circ\text{C}^{-1}$, of the material of which the bottle is made (glass or plastic), and $\Delta\theta \text{ } ^\circ\text{C}$ is the difference in temperature between the inside and the outside of the bottle. Put

$$\kappa \left\{ \frac{2\pi l}{\log_e((r+b)/r)} + 2\frac{\pi r^2}{b} \right\} = K$$

to save writing. Then in a short time δt the liquid gains an amount of heat energy

$$q\delta t = K(\theta_s - \theta)\delta t,$$

where $\theta_s \text{ } ^\circ\text{C}$ is the temperature of the surroundings. The resulting increase in the temperature of the liquid, $\delta\theta \text{ } ^\circ\text{C}$, is given approximately by

$$\delta\theta = \frac{q\delta t}{mc} = \frac{K}{mc}(\theta_s - \theta)\delta t,$$

where $m \text{ kg}$ is the mass of the liquid and $c \text{ J kg}^{-1} \text{ } ^\circ\text{C}^{-1}$ its specific heat. On letting δt tend to zero we obtain

$$\frac{d\theta}{dt} = \frac{K}{mc}(\theta_s - \theta).$$

This differential equation is the required model of the warming process.

If we put $\phi = \theta - \theta_s$ and $\lambda = K/(mc)$, the equation becomes

$$\frac{d\phi}{dt} = -\lambda\phi,$$

as in the text. The solution is

$$\phi = \phi_0 e^{-\lambda t},$$

or

$$\theta(t) = \theta_s + (\theta_0 - \theta_s)e^{-\lambda t},$$

where θ_0 is the initial temperature of the liquid. Note that since $\theta_0 < \theta_s$, the temperature of the liquid is in fact increasing and tending to θ_s from below: perhaps it is better to write the solution as

$$\theta(t) = \theta_s - (\theta_s - \theta_0)e^{-\lambda t},$$

to make this a little more obvious.

Suppose that the temperature at which the liquid becomes unpleasant to drink is θ_u °C. The time it takes for the liquid to reach this temperature is then

$$\frac{1}{\lambda} \log_e \left(\frac{\theta_s - \theta_0}{\theta_s - \theta_u} \right).$$

To find a numerical value, the values of the various parameters need to be known.

The results of this analysis are very similar in appearance to those of the model of the cooling of tea considered in the text, even though that is a problem of cooling by convection while this is a problem of warming by conduction. In fact, any situation involving heat transfer processes which, in the steady state, lead to a rate of heat transfer which is proportional to temperature difference, will give a similar answer. Of course, the nature of the heat transfer processes, the shape of the container and so on will affect the outcome – but they will do so only through the value of λ . This observation suggests a different approach to the evaluation of parameters. One could ignore the details of the process, and find the value of λ directly by experiment instead. Thus to deal with the warming of a bottle of milk, it would be enough to conduct one experiment to find the value of λ for a milk bottle; this could be done by noting that the time taken for the temperature difference between the contents of the bottle and the surroundings to halve is $(\log_e 2)/\lambda$. With the value of λ known, the formula derived above will give the time it takes for a bottle of milk to reach any given temperature when taken from the refrigerator, and it can also be used to determine what the temperature of the milk will be a given time after it has been collected from the doorstep and put into the refrigerator.

2. The temperature inside the house while it is cooling (with the heating off) obeys the law obtained in the text and the previous exercise, namely

$$\theta(t) = \theta_s + (\theta_0 - \theta_s)e^{-\lambda t},$$

where now $\theta(t)$ °C is the temperature in the house t seconds after the thermostat has turned the heating off, θ_s °C is the temperature outside the house (assumed constant), and θ_0 °C is the temperature at which the heating cuts out (called the ‘predetermined higher value’ in the question).

Here λ is a parameter which measures the rate of heat loss for the whole house, which could be defined as $(\log_e 2)/T$ where T is the time taken for the temperature difference between the inside and the outside of the house to fall to half its initial value: see the solution to the previous exercise.

The heating remains off until the temperature inside the house falls to the ‘predetermined lower value’, at which point the heating comes on. By assumption, the heating system supplies heat at a constant rate, say H watts. But while the heating is on, the house continues to lose heat by the same processes as before, as well as gaining heat from the heating system. We must therefore modify the model of

cooling to take account of the heat supplied. Formerly, we had the expression

$$q\delta t = K(\theta_s - \theta)\delta t,$$

for the amount of heat transferred in time δt ; here $K = \lambda mc$, where now m kg is the mass and c J kg⁻¹ °C⁻¹ the specific heat of the air in the house. To take account of the heat supplied we must change this to

$$q\delta t = \{K(\theta_s - \theta) + H\}\delta t.$$

The differential equation for the temperature inside the house now becomes

$$\frac{d\theta}{dt} = \lambda(\theta_s - \theta) + \frac{H}{mc}.$$

Before rushing in to solve this equation as it stands, let us note that it predicts that $d\theta/dt = 0$ when

$$\theta = \theta_s + \frac{H}{\lambda mc}.$$

Thus the temperature of the house can stay constant at this value, with the heating on. We call this temperature the equilibrium temperature and denote it by θ_e . At the equilibrium temperature the heat supplied just balances the heat lost to the outside. In terms of θ_e the differential equation can be rewritten as

$$\frac{d\theta}{dt} = \lambda(\theta_e - \theta).$$

The solution has the same form as before, but with θ_e replacing θ_s : so in the heating phase we have

$$\theta(t) = \theta_e + (\theta_1 - \theta_e)e^{-\lambda t},$$

where now $\theta(t)$ °C is the temperature in the house t seconds after the thermostat has turned the heating on, θ_e °C is the equilibrium temperature when the heating is on, and θ_1 °C is the temperature at which the heating cuts in (called the ‘predetermined lower value’ in the question). We must assume that the heating system is sufficiently powerful that $\theta_e > \theta_0$, since otherwise the temperature inside the house will never reach the value at which the heating cuts out. Assuming this to be so, the heating will continue to run until the temperature θ_0 is reached, when the thermostat will operate again, and the cooling half of the cycle will recommence. The temperature in the house will vary as shown in Figure 1.

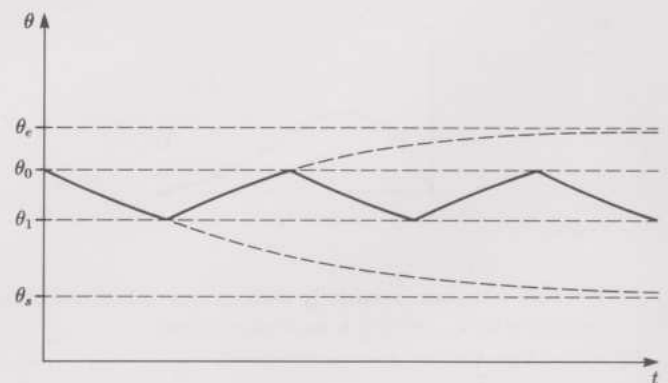


Figure 1

The difference between the equilibrium temperature θ_e and the external temperature θ_s depends just on the rate of supply of heat and the thermal characteristics of the house; so when the external temperature drops, the equilibrium temperature drops by the same amount. But the thermostat temperatures θ_0 and θ_1 do not change. So when it is colder outside, the cooling and heating curves move downwards

relative to θ_0 and θ_1 , with the result that the cooling curve is steeper between θ_0 and θ_1 , and the heating curve less steep. Thus when it is colder outside, the heating part of the cycle is longer and the cooling part shorter, as one would expect.

This problem has some features in common with the drug therapy model: the range of temperatures allowed by the thermostat is like the therapeutic range for the concentration of the drug; while the heating part of the cycle and the intravenous infusion method of administering a drug are modelled in very similar ways. Models occurring in apparently totally different circumstances can be surprisingly similar; it can often be helpful when modelling something unfamiliar to ask yourself 'have I ever come across anything similar before?'

Solution to the exercise in Section 5

1. (i) Things like convenience of handling and stacking must, presumably, be taken into account; but surely the most important thing, from the manufacturer's point of view, is to use as little tin-plate as possible consistent with the fact that the can has to hold a given amount of beans (or whatever foodstuff it is destined for).

(ii) The problem is: to find a formula for the area of tin-plate required to make a cylindrical can which will contain a specified volume of food; this formula should give the area in terms of the dimensions of the can (its height and the radius of either end); and then to use this formula to find the dimensions of the can for which the area of tin-plate used is a minimum.

It seems that canned food is sold by volume rather than weight (or more properly mass). The evidence for this comes from my kitchen cupboard, where I find cans of different foods the masses of whose contents range from 397 g to 432 g; the cans are identical in size and shape. It is clearly better to have a standard size and shape of can, and therefore a standard volume, rather than slightly different sizes of can for 400 g of peas, 400 g of beans and 400 g of pears (for example).

(iii) The fundamental variables are the area A of tin-plate used in the manufacture of a can, the height h of the can, and the radius of an end, r . In SI units the linear dimensions would be in metres; but in this case the metre is rather too large a unit, the millimetre rather too small, and centimetres would seem to be the better choice. Then A must be measured in cm^2 . Of course, one could use inches and square inches.

(iv) The volume is a parameter rather than a variable. The fact that we are concerned only with cans which enclose a given volume will enable us to formulate a relation which must be satisfied by the dimensions of the can, which will enable us to eliminate one of them (either the height of the can or the radius of an end) and express the area entirely in terms of the other.

(v) The can is made up of one rectangle of sides h and $2\pi r$ (which forms the curved body of the can) and two circular ends, each of area πr^2 . Thus

$$A = 2\pi r h + 2\pi r^2.$$

The volume V which the can encloses is

$$V = \pi r^2 h.$$

Eliminating h gives

$$A = 2\pi r \left(\frac{V}{\pi r^2} \right) + 2\pi r^2 = 2 \left(\frac{V}{r} + \pi r^2 \right).$$

(vi) The area of tin-plate required to make the seams has been ignored. So also has the waste resulting from cutting circular pieces for the ends of the can from sheet material.

(vii) We have $A \rightarrow \infty$ as $r \rightarrow 0$ and as $r \rightarrow \infty$. In the former case the can becomes long and thin; in the latter, short and fat.

(viii) Since the area is always positive, it must have at least one local minimum. In fact, it is easy to sketch the graph of A by noting that it is the sum of a quadratic function and a multiple of $1/r$: the graph looks like this.

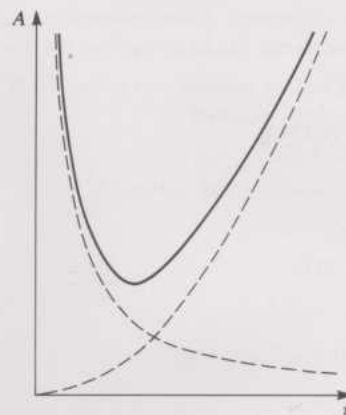


Figure 1

There would appear to be just one local minimum, which will provide the solution to the problem.

We have

$$A = 2 \left(\frac{V}{r} + \pi r^2 \right),$$

where V is a constant. Thus

$$\frac{dA}{dr} = 2 \left(-\frac{V}{r^2} + 2\pi r \right),$$

and so $dA/dr = 0$ where

$$r^3 = \frac{V}{2\pi}.$$

It is clear from the graph that this is a minimum; but to confirm, note that

$$\frac{d^2A}{dr^2} = 4 \left(\frac{V}{r^3} + \pi \right),$$

which is positive for all positive r .

The ratio of height to radius of a can of minimum area is given by

$$\frac{h}{r} = \frac{V/(\pi r^2)}{r} = \frac{V}{\pi r^3} = 2.$$

This means that for a can of minimum area, the height is equal to the diameter of the base; in other words, seen from the side the can will appear to be square. In fact, most cans are taller than they are wide.

(ix) The most significant source of error is the fact that we have ignored the wastage in cutting the circular ends; in computing the area of tin-plate used to make the can, we should include the waste as well as the tin-plate which actually ends up as the can.

(x) The simplest assumption is that each end is cut from a square of material of side $2r$, so that the formula for the area becomes

$$A = \frac{2V}{r} + 8r^2.$$

The minimum now occurs where

$$-\frac{2V}{r^2} + 16r = 0,$$

or

$$r^3 = \frac{V}{8}.$$

The ratio of height to radius becomes

$$\frac{h}{r} = \frac{V}{\pi r^3} = \frac{8}{\pi} \simeq 2.55.$$

This gives a shape more like that of actual cans, but still rather too short and fat: the corresponding ratio for an actual can is 3.10. We still have to take the seams into account, but it is hard to believe that this will make a substantial difference. This analysis suggests that there is money to be saved in changing the shape of tin cans.

(xi) For a square section can of height h and width w the area of tin-plate required is

$$A = 4hw + 2w^2$$

(there is no waste). The volume is

$$V = hw^2,$$

so we may write

$$A = \frac{4V}{w} + 2w^2.$$

The minimum occurs where

$$-\frac{4V}{w^2} + 4w = 0,$$

or

$$w = V^{1/3} = h;$$

the optimal shape is a cube (as one might have guessed), and the area required is $6V^{2/3}$. For the solution found in the previous part we have $r = \frac{1}{2}V^{1/3}$, so $A = 4V^{2/3} + 2V^{2/3} = 6V^{2/3}$ also. Thus square section cans use the same area of tin-plate for a given volume.

